

Supplemental document to

Early Pleistocene enamel proteome sequences from Dmanisi resolve *Stephanorhinus* phylogeny

General remark.

This document contains a selection of MS/MS spectra, and the related b and y fragment ion series, supporting the identification of the peptides covering the positions of phylogenetic interest (SAPs) in the enamel proteome of *Stephanorhinus* specimen Dm.5/157-16635. Such positions are marked by gray bars, with amino acid coordinates provided at the top. Note that for each position many other peptides/spectrum matches in either MaxQuant or PEAKS were observed (data not shown).

For each tandem mass spectrum we also provide the "stagetip number - scan number" for easy referral to the raw data. For several spectra we also confirmed the peptide identification through manual annotation and/or synthetic peptides. The manual and automatically annotated spectra are displayed on the same page (top=MaxQuant, bottom=manual/synthetic).

Synthetic peptides and their corresponding ancient spectra are shown at the end of the document. Annotation of the synthetic and ancient peptides is performed manually through the Interactive Peptide Spectral Annotator freeware.

For manual annotation, annotation abbreviations are as follows:

y/b⁰ = y/b-H₂O

y/b^{*} = y/b-NH₃

y^Δ = y-H₃PO₄

y^{ΔΔ} = y-2H₃PO₄

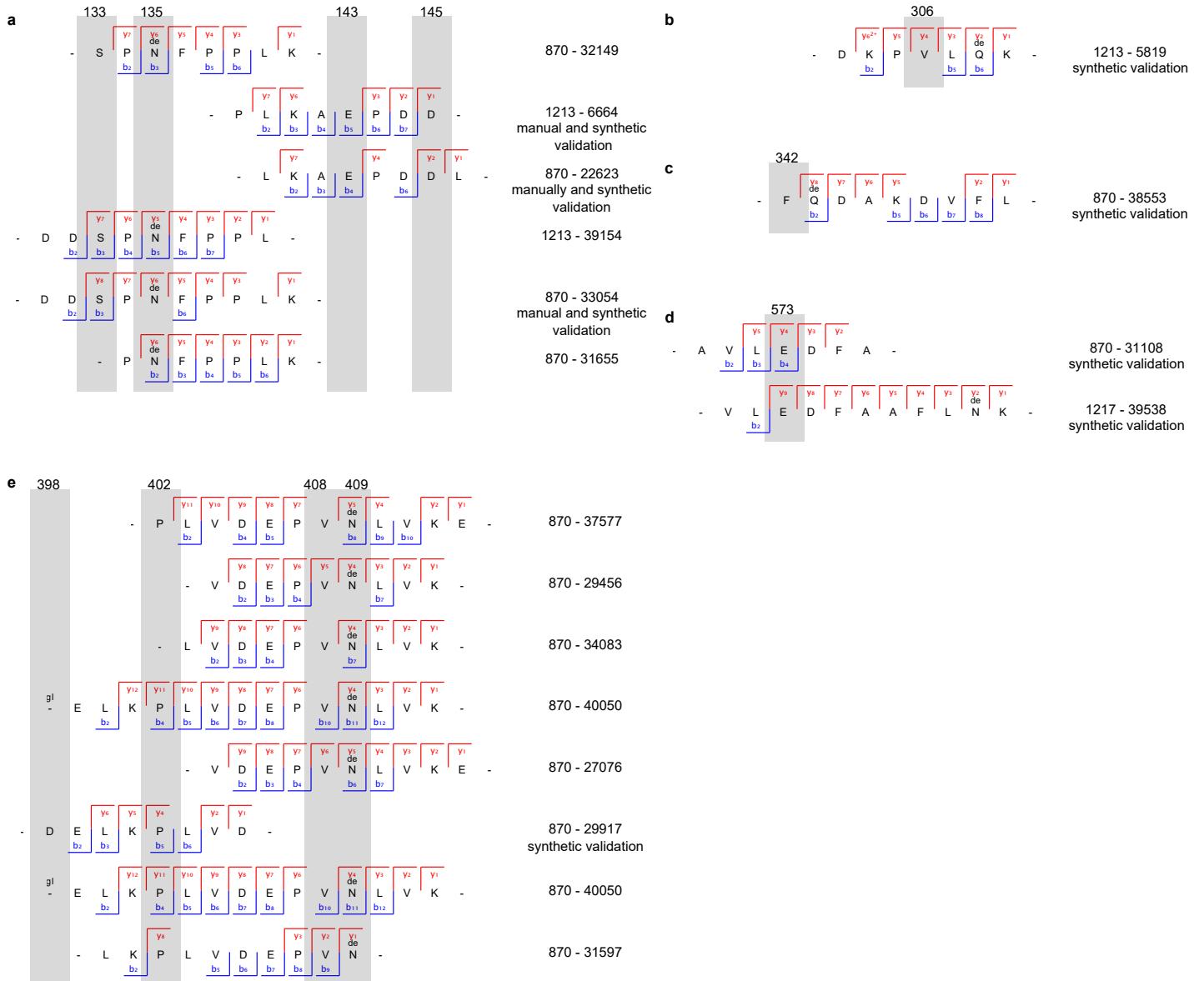
For MQ annotation, annotation abbreviations including a * indicate the presence of a neutral phosphate loss on the fragment ion (-98Da, -H₃PO₄).

Several spectra, including two manually verified, indicate the presence of endogenous phosphorylation.

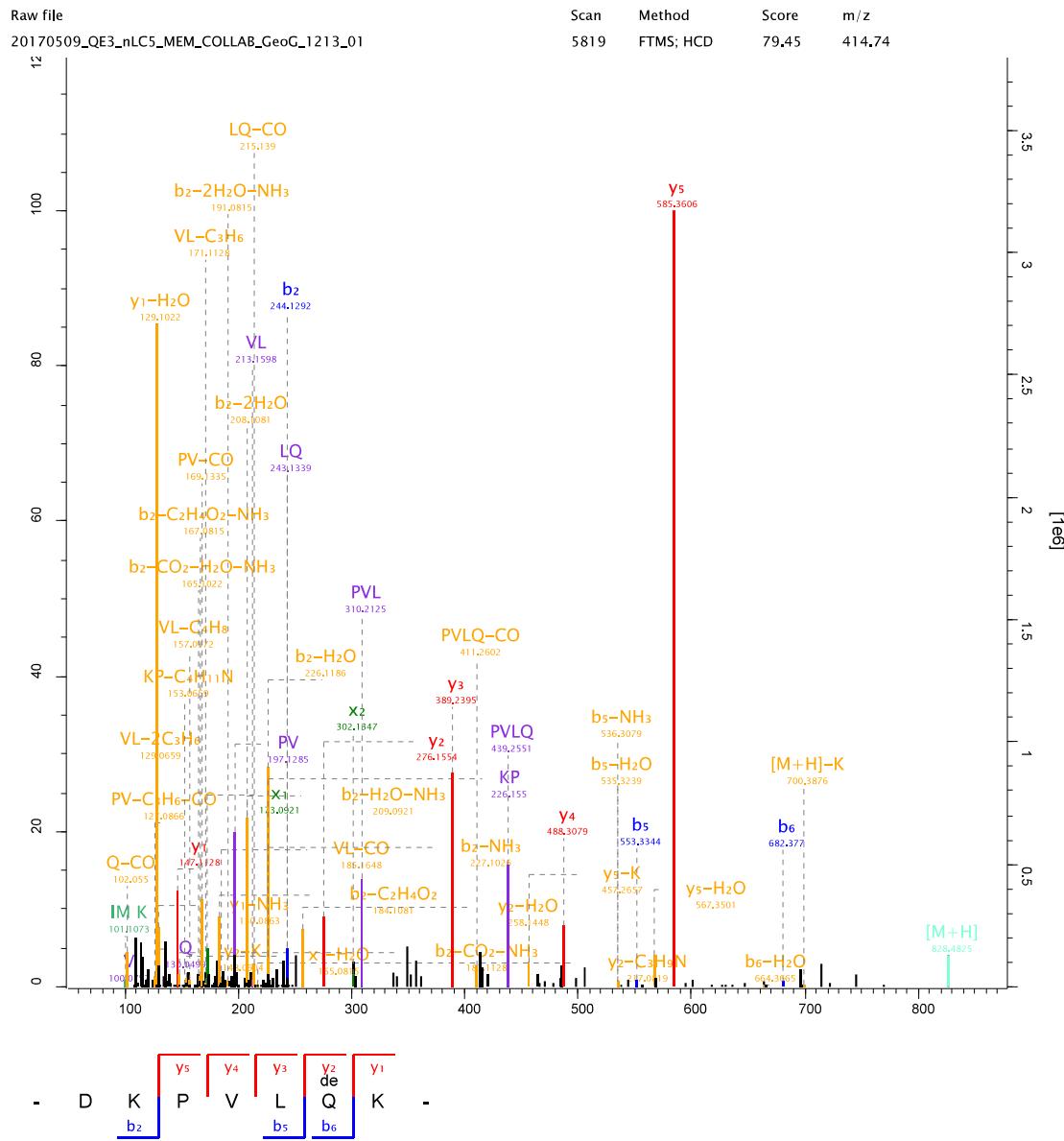
Leucines and isoleucines cannot be differentiated from each other in our MS set-up.

(Stagetip - scan number)

(Stagetip - scan number)



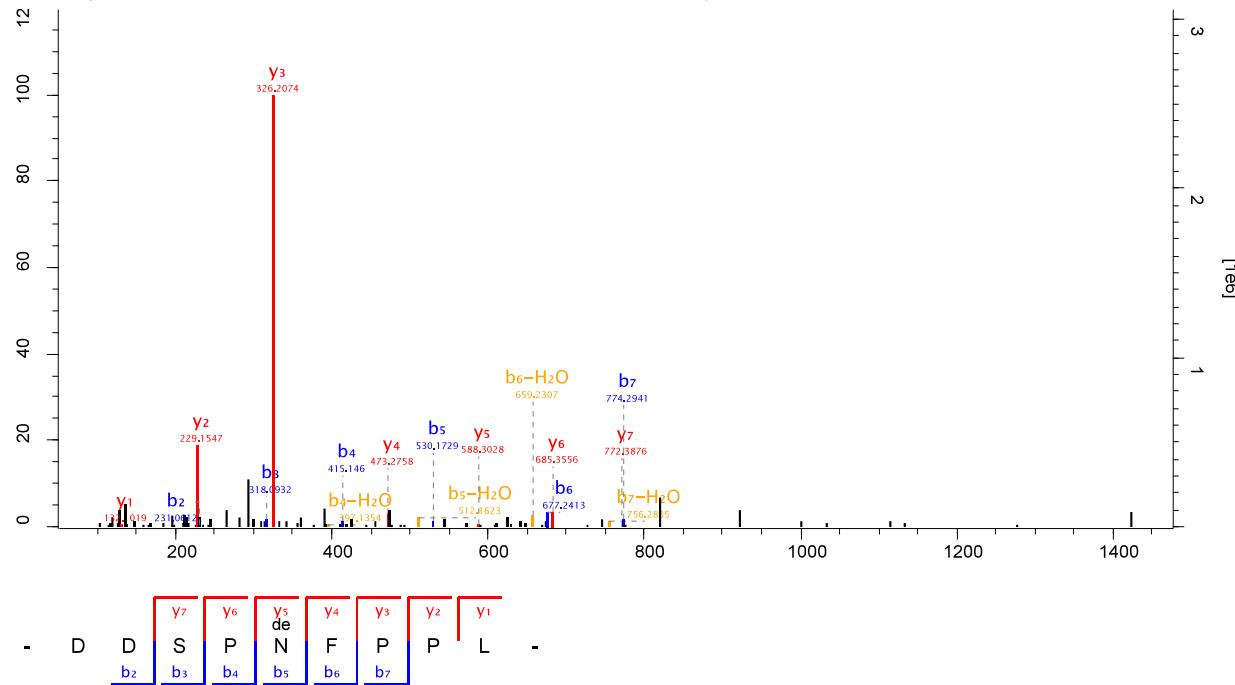
ALB. Peptide matches to phylogenetically informative positions 133, 135, 143, 145 (a), 306 (b), 342 (c), 573 (d), and 398, 402, 408, 409 (e). Coordinates in reference to XP_004419220.

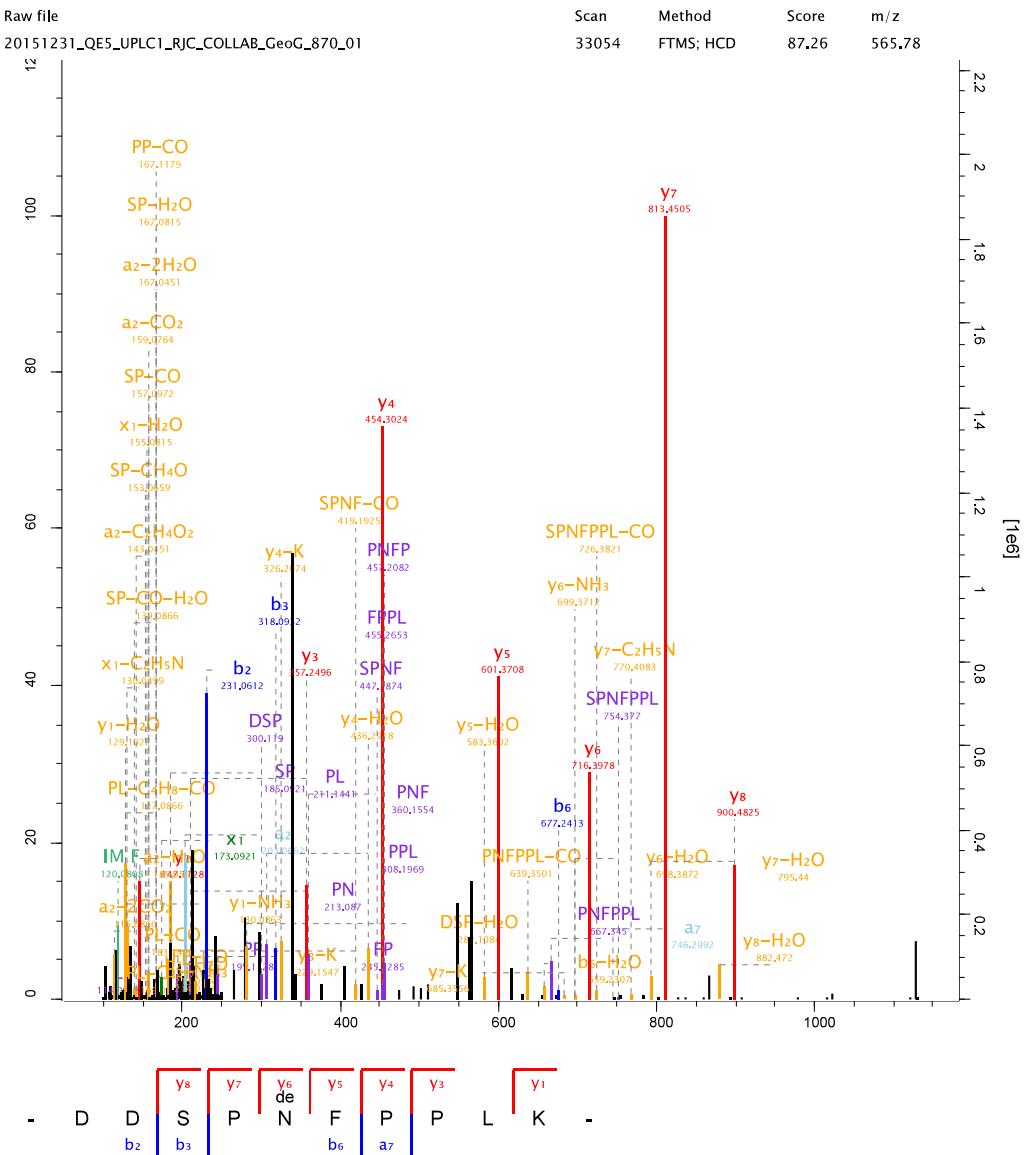


- D K P V L Q K -
 b₂ Y₅ Y₄ Y₃ Y₂ de b₅ Y₁

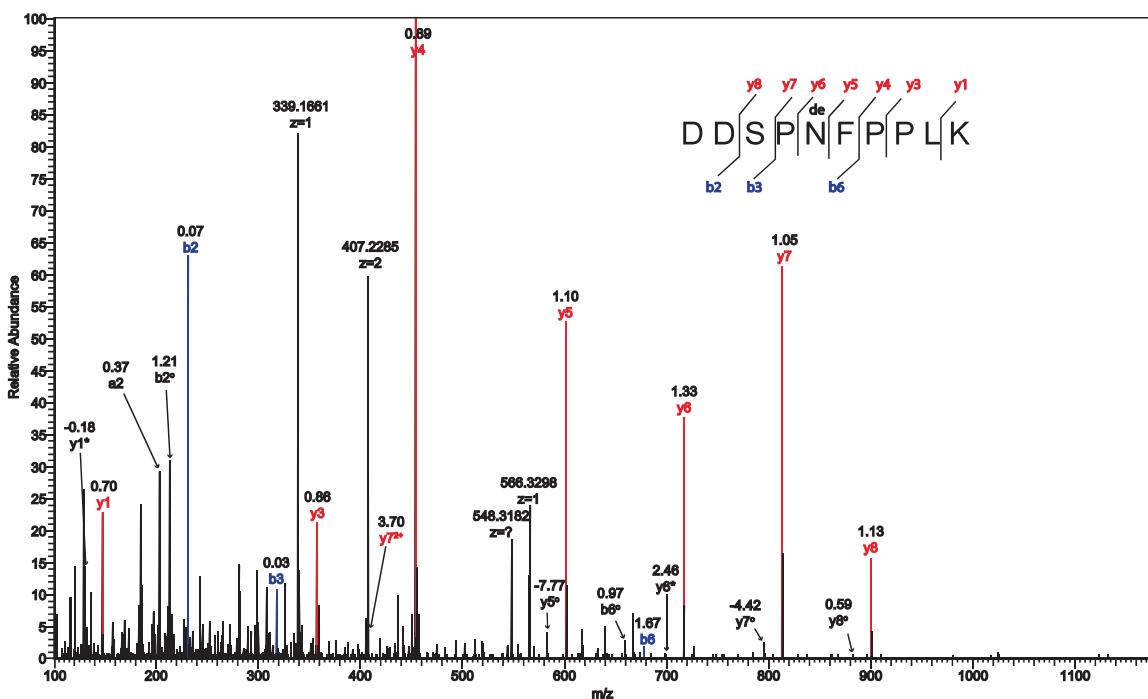
Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1213_01

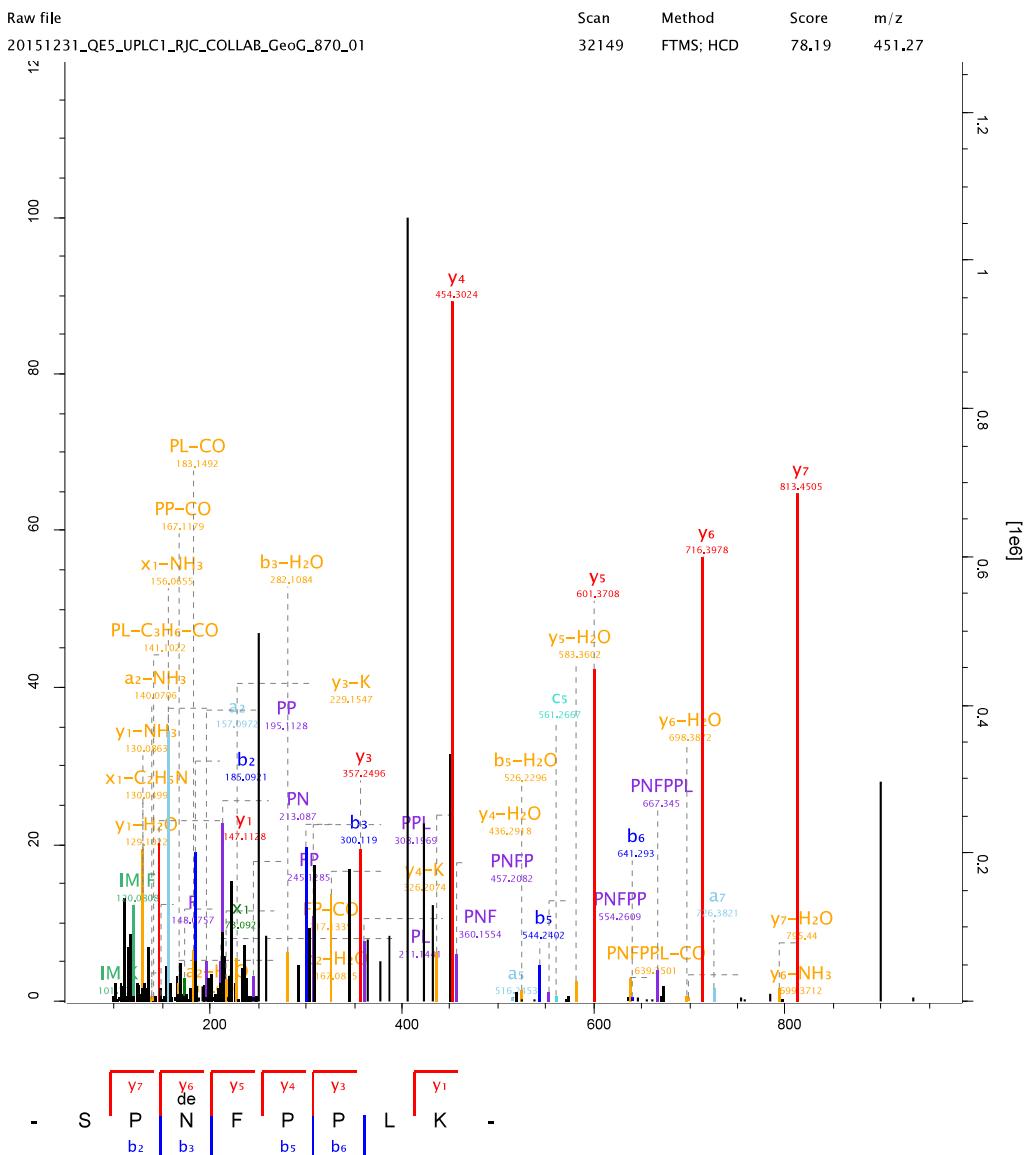
Scan 39154 Method FTMS; HCD Score 77.07 m/z 1002.44



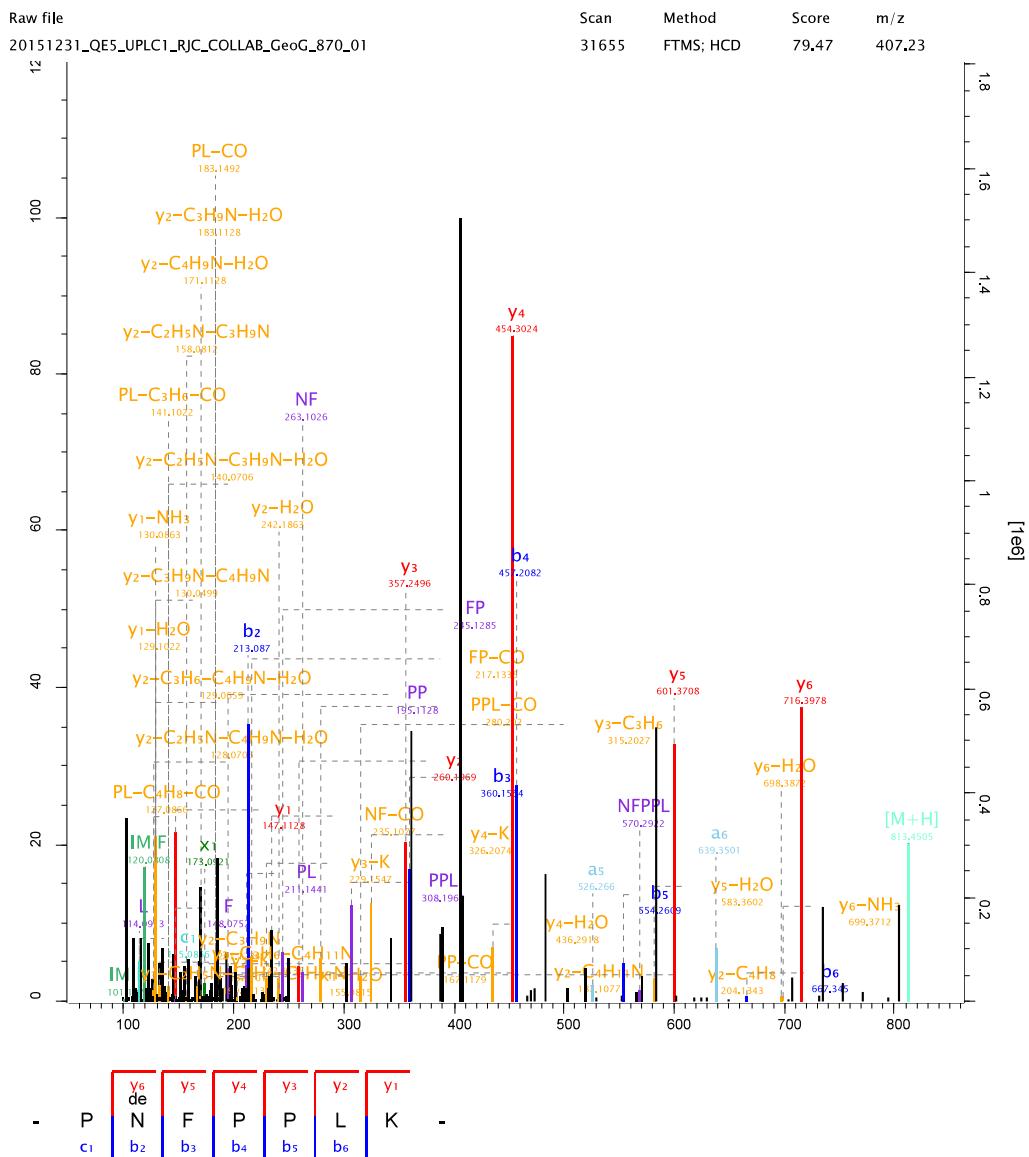


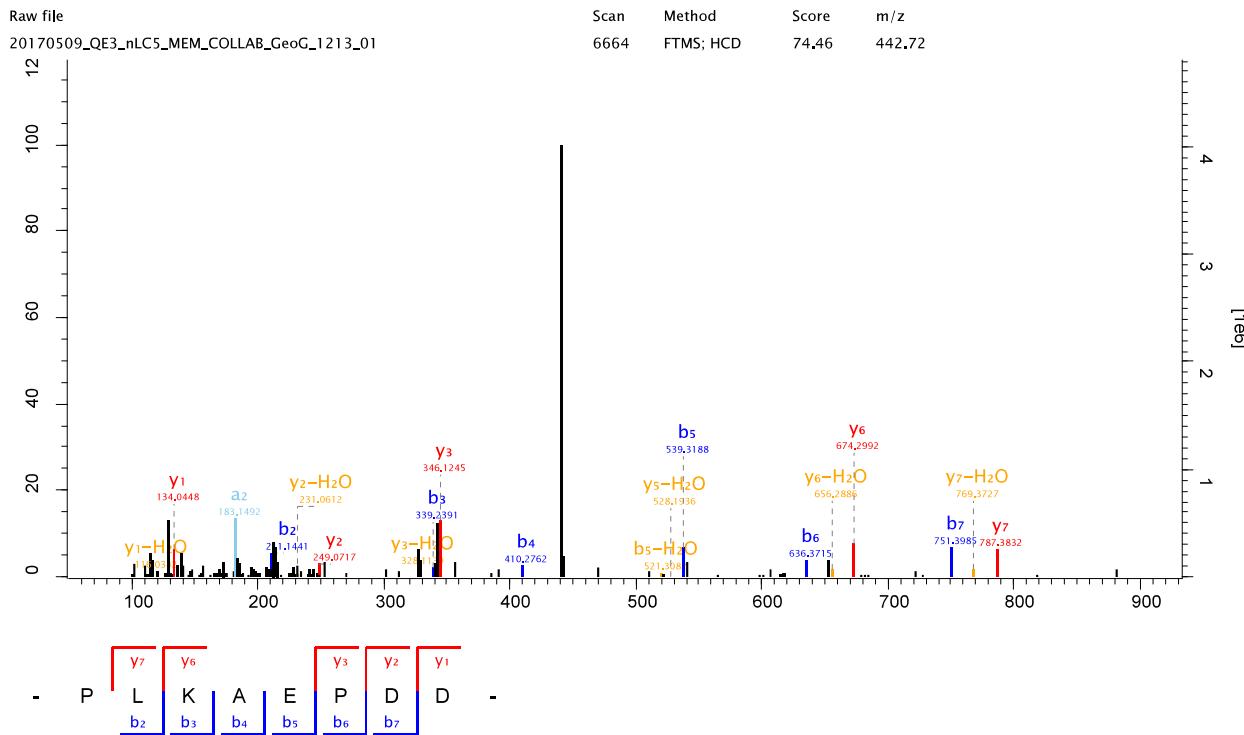
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01 #33054 RT: 109.53 AV: 1 NL: 2.97E5
T: FTMS + p NS1d Full ms2 565.78@hcd28.00 [100.00-1175.00]



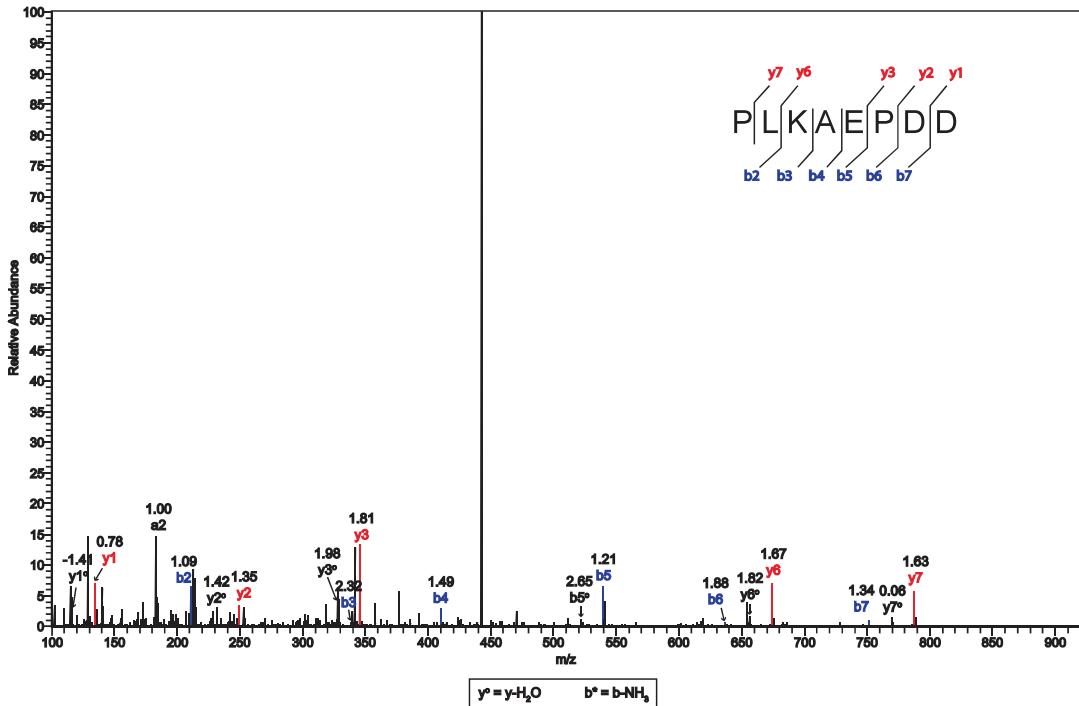


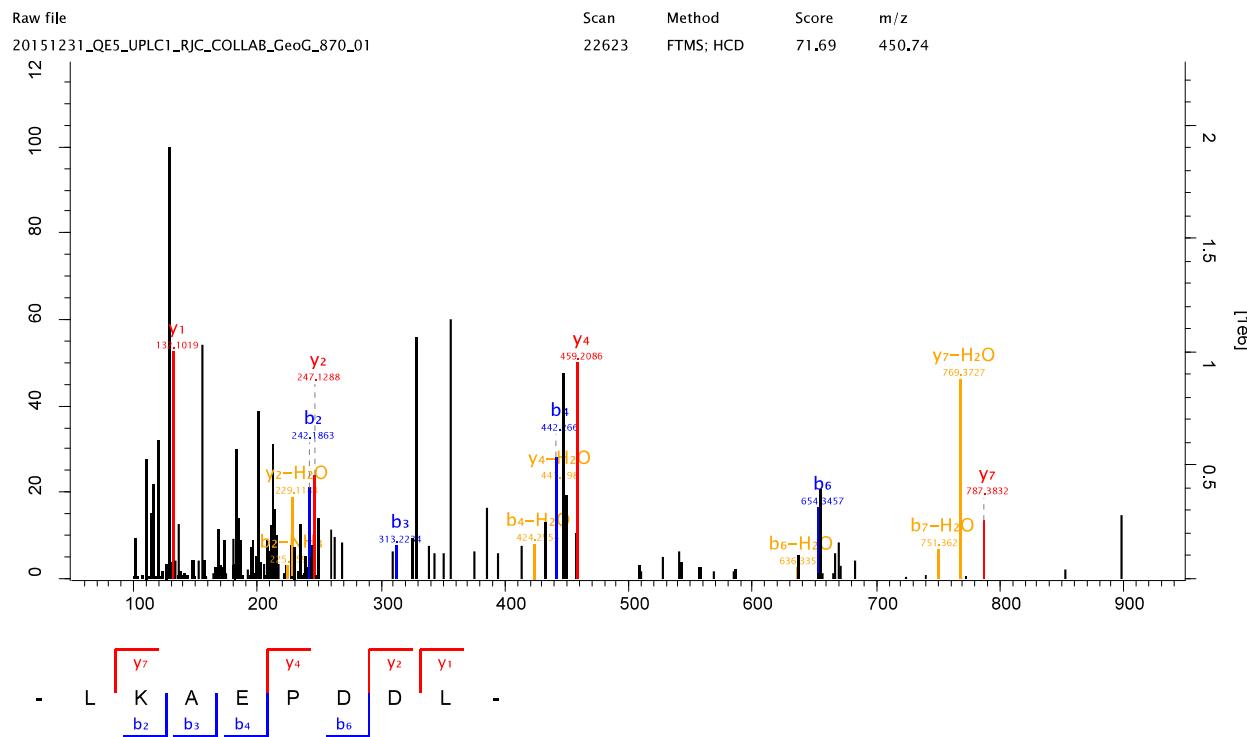
S P b₂ y₇ de N b₃ F y₆ y₅ P y₄ y₃ P b₅ L y₁ K



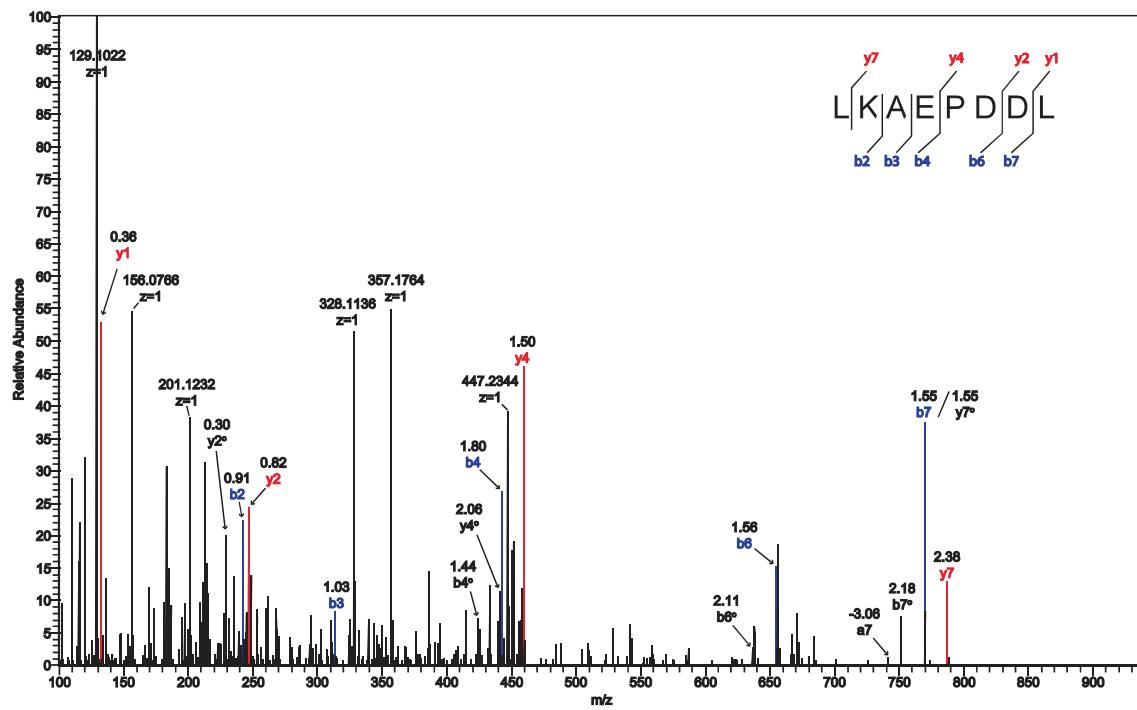


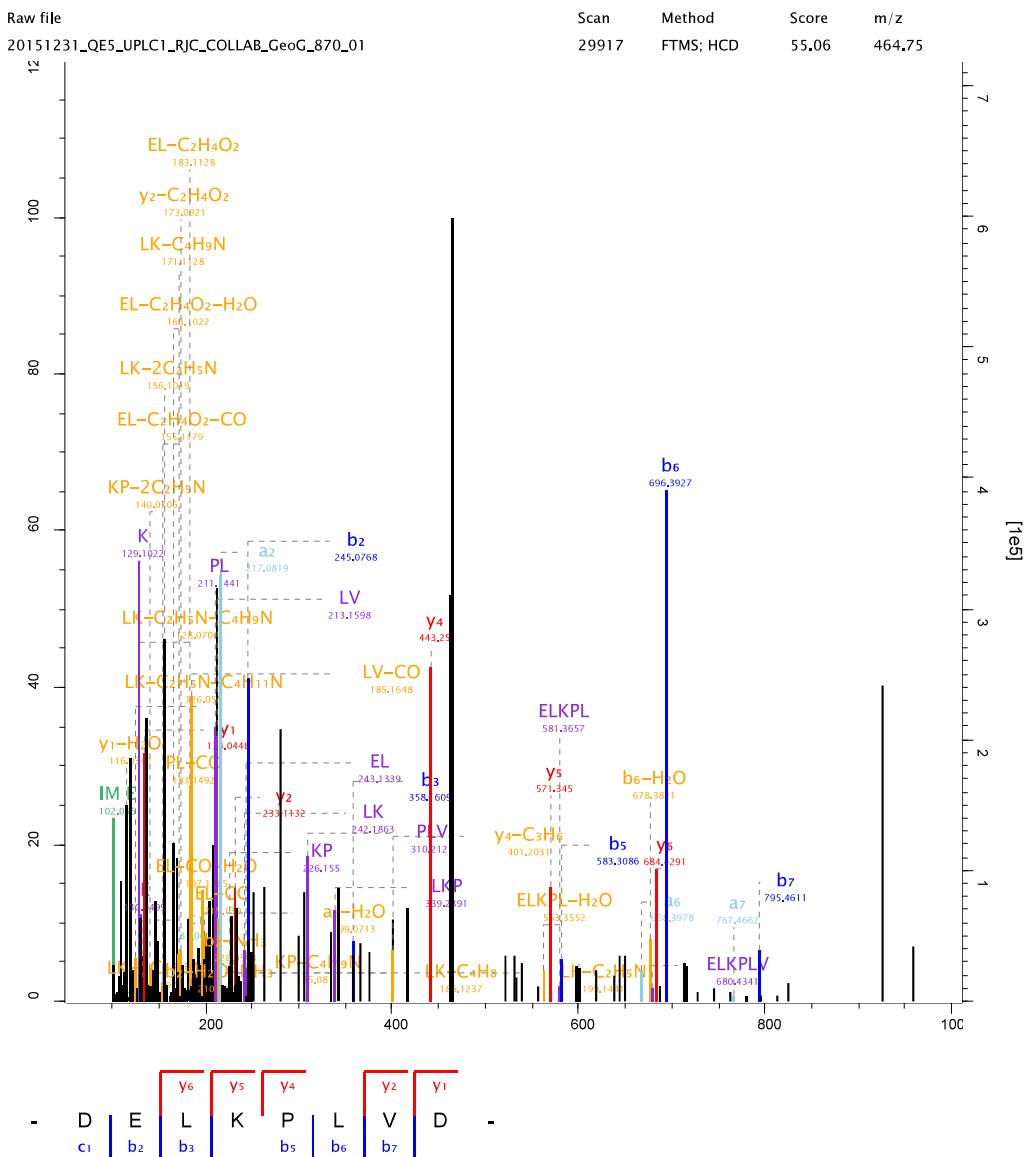
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1213_01 #6664 RT:23.91 AV:1 NL:9.06E5
T: FTMS +p NSId Full ms2 442.72@hcd28.00 [700.00-920.00]

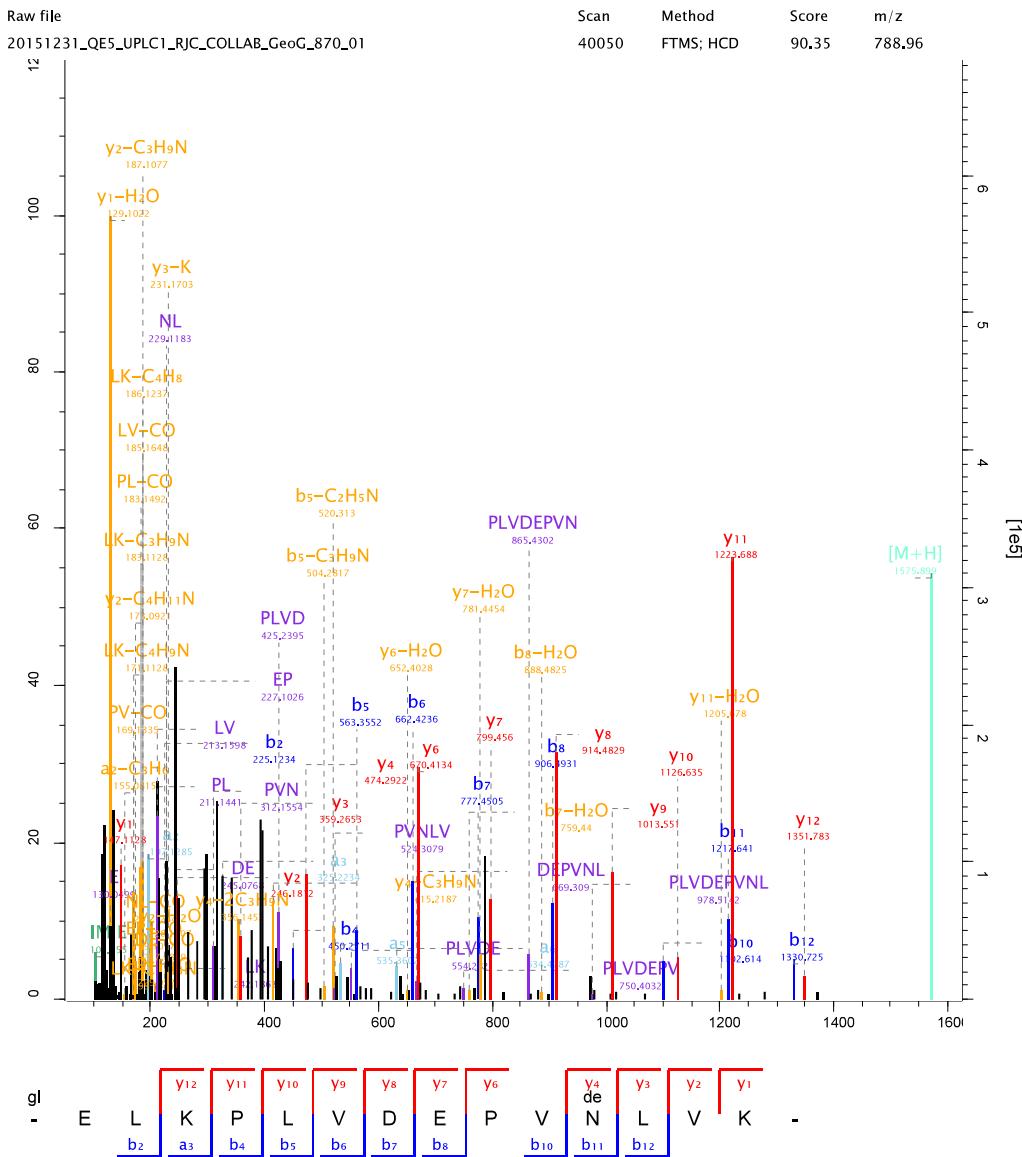




20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01 #22623 RT: 77.15 AV: 1 NL: 4.70E5
T: FTMS + p NSId Full ms2 450.74@hcd28.00 [100.00-940.00]



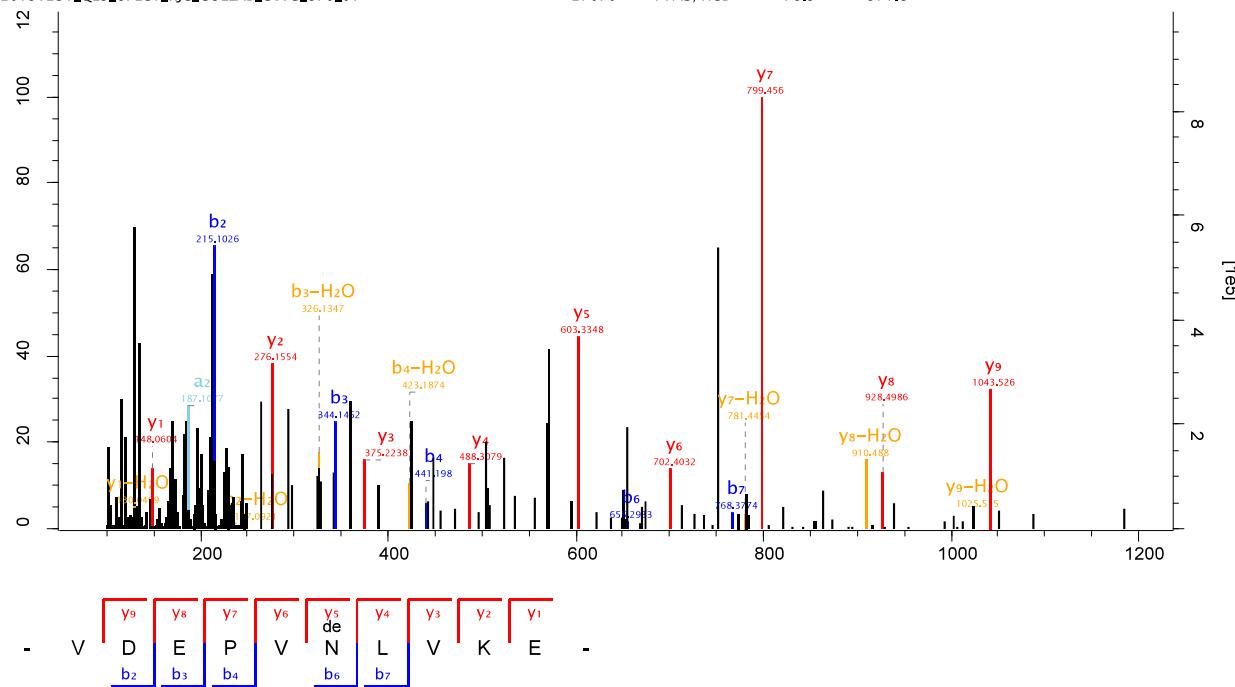


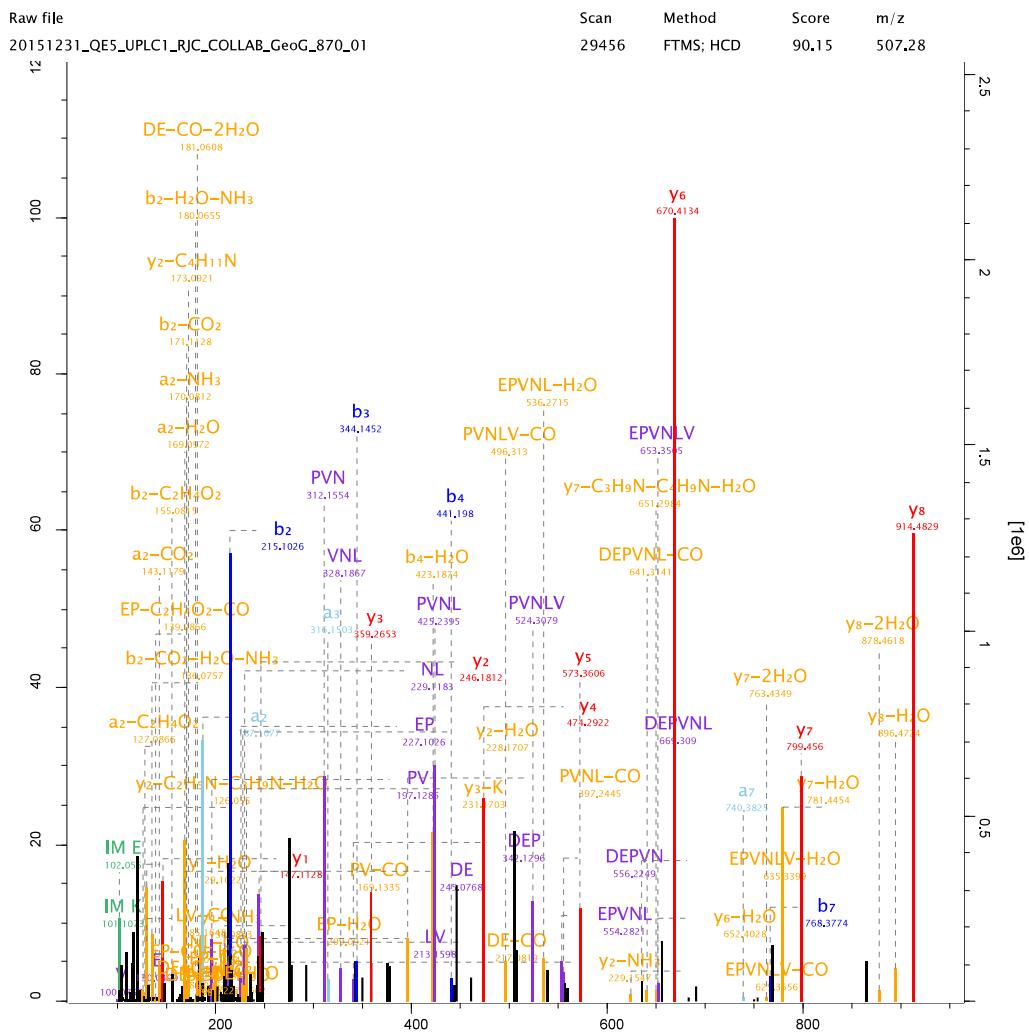


gl	E	y_{12}	y_{11}	y_{10}	y_9	y_8	y_7	y_6	y_5	y_4	y_3	y_2	y_1
-	L	K	P	L	V	D	E	P	V	N	L	V	K
-	b_2	a_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	-	-

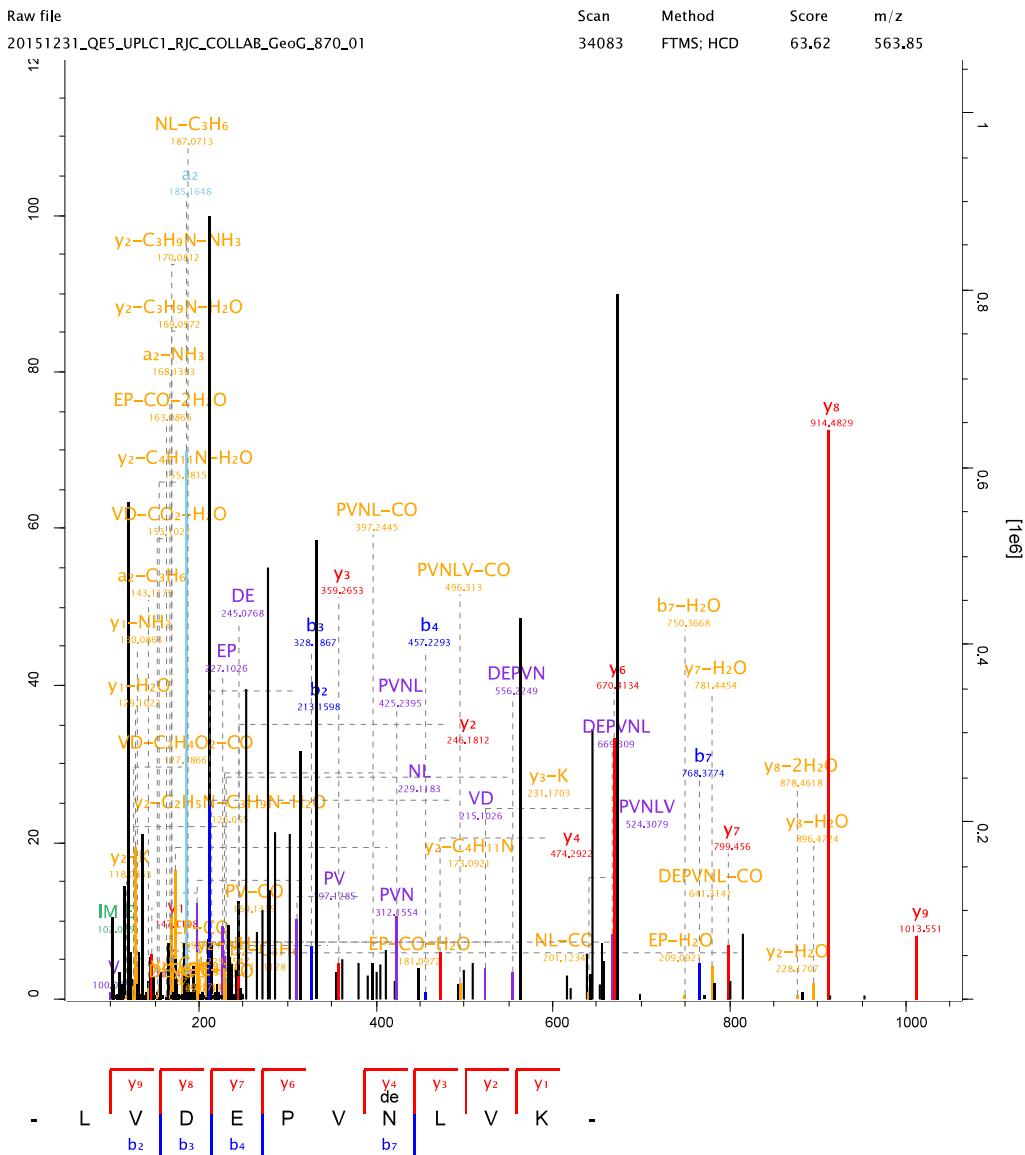
Raw file
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01

Scan 27076 Method FTMS; HCD Score 78.9 m/z 571.8





A morphological tree diagram for the word 'de Vlaams'. The root node is 'V' (Verb). It branches into 'de' (dative case marker) and 'Vlaams' (adjective). 'de' further branches into 'de' (definite article) and 'Vlaams' (adjective). 'Vlaams' branches into 'Vlaam' (stem) and 's' (singular suffix).



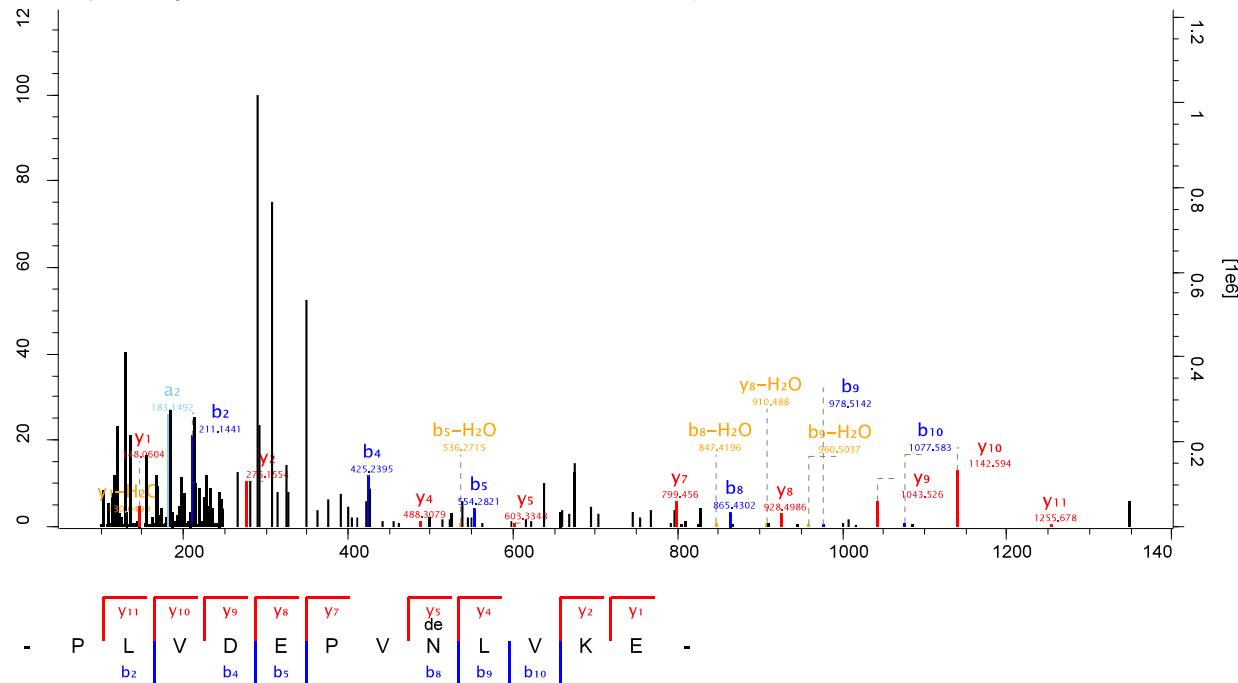
L V D E P V N L V K

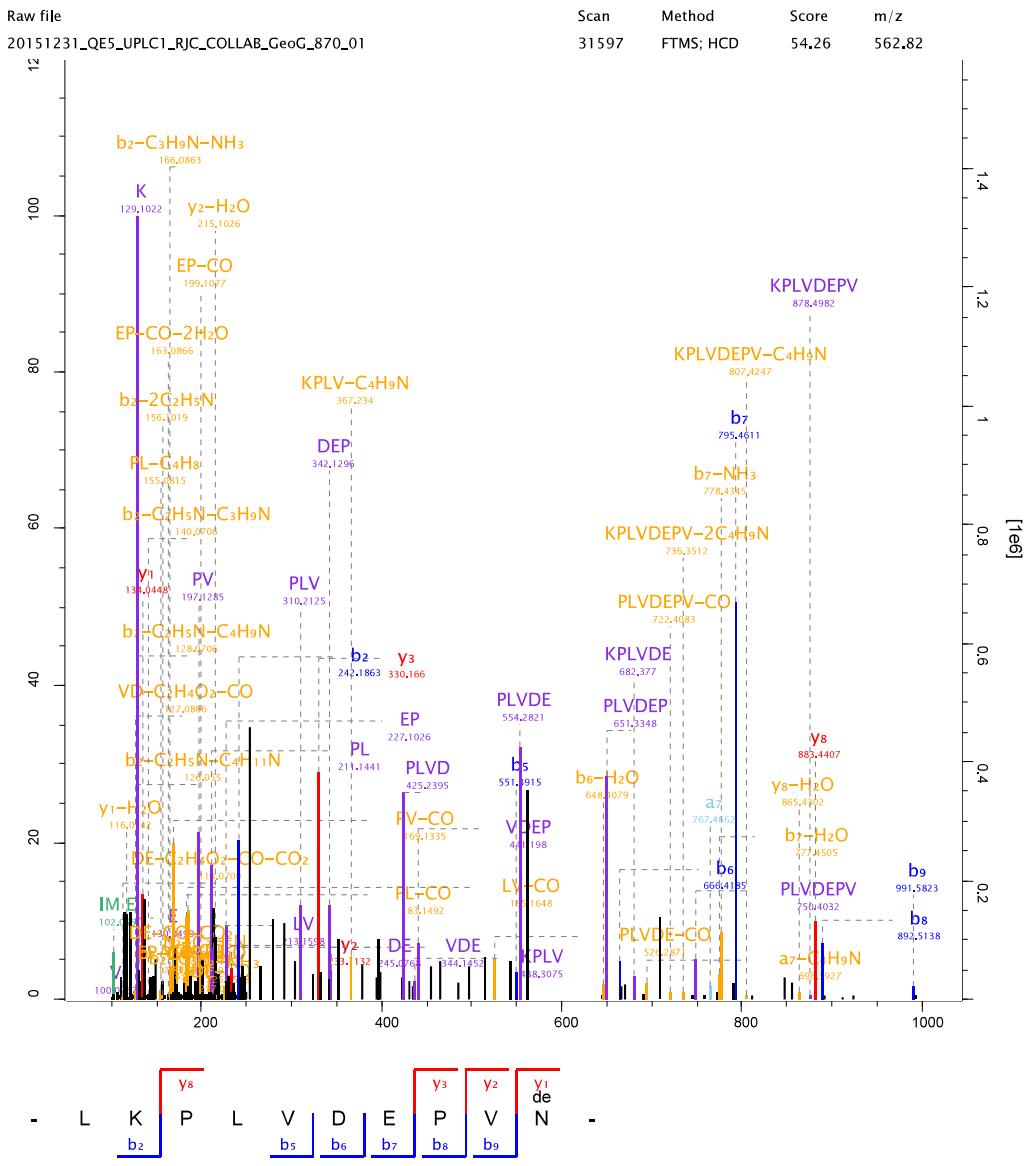
Y₉ Y₈ Y₇ Y₆ de Y₃ Y₂ Y₁

b₂ b₃ b₄ b₇

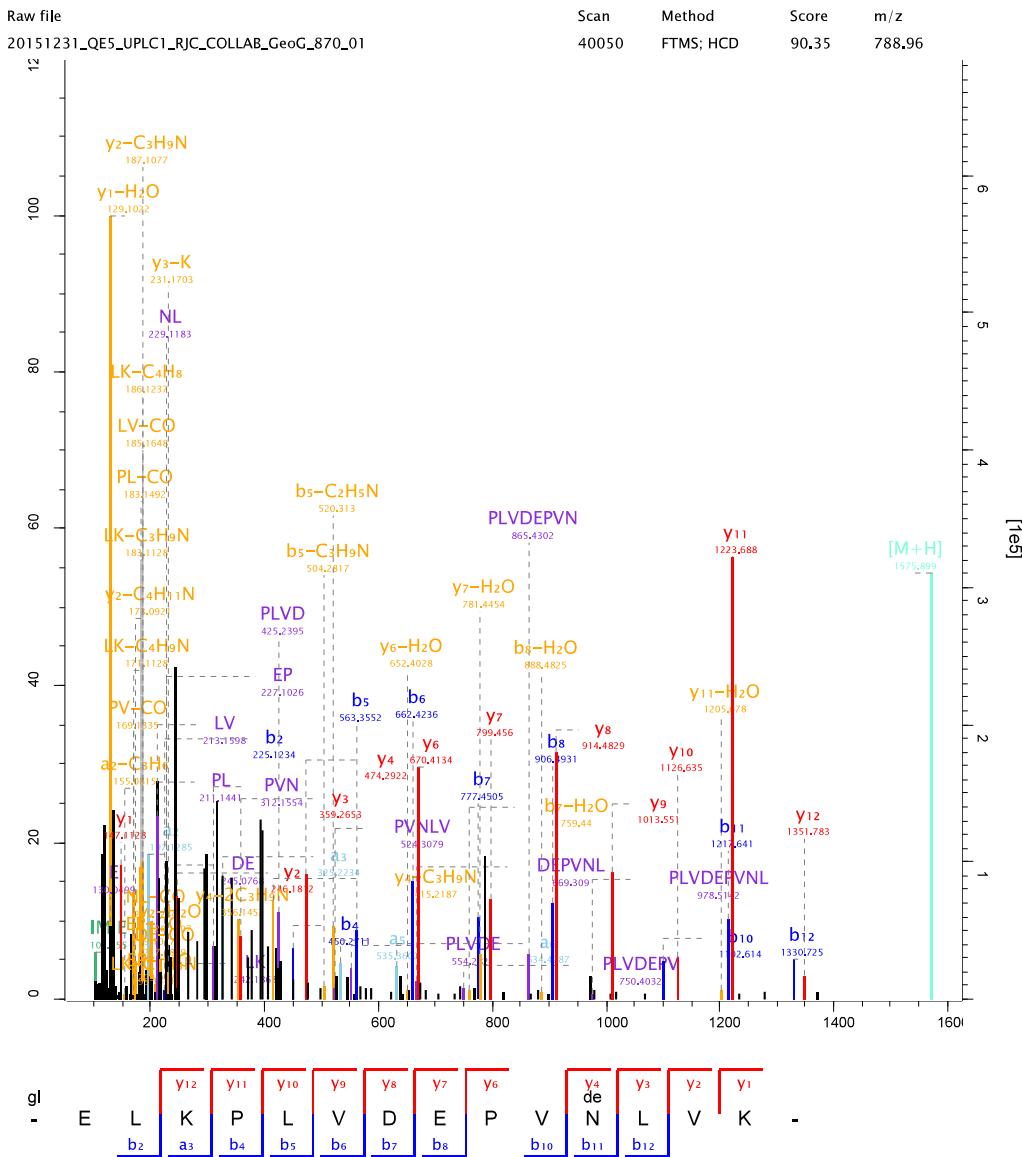
Raw file
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01

Scan 37577 Method FTMS; HCD Score 51.01 m/z 676.87

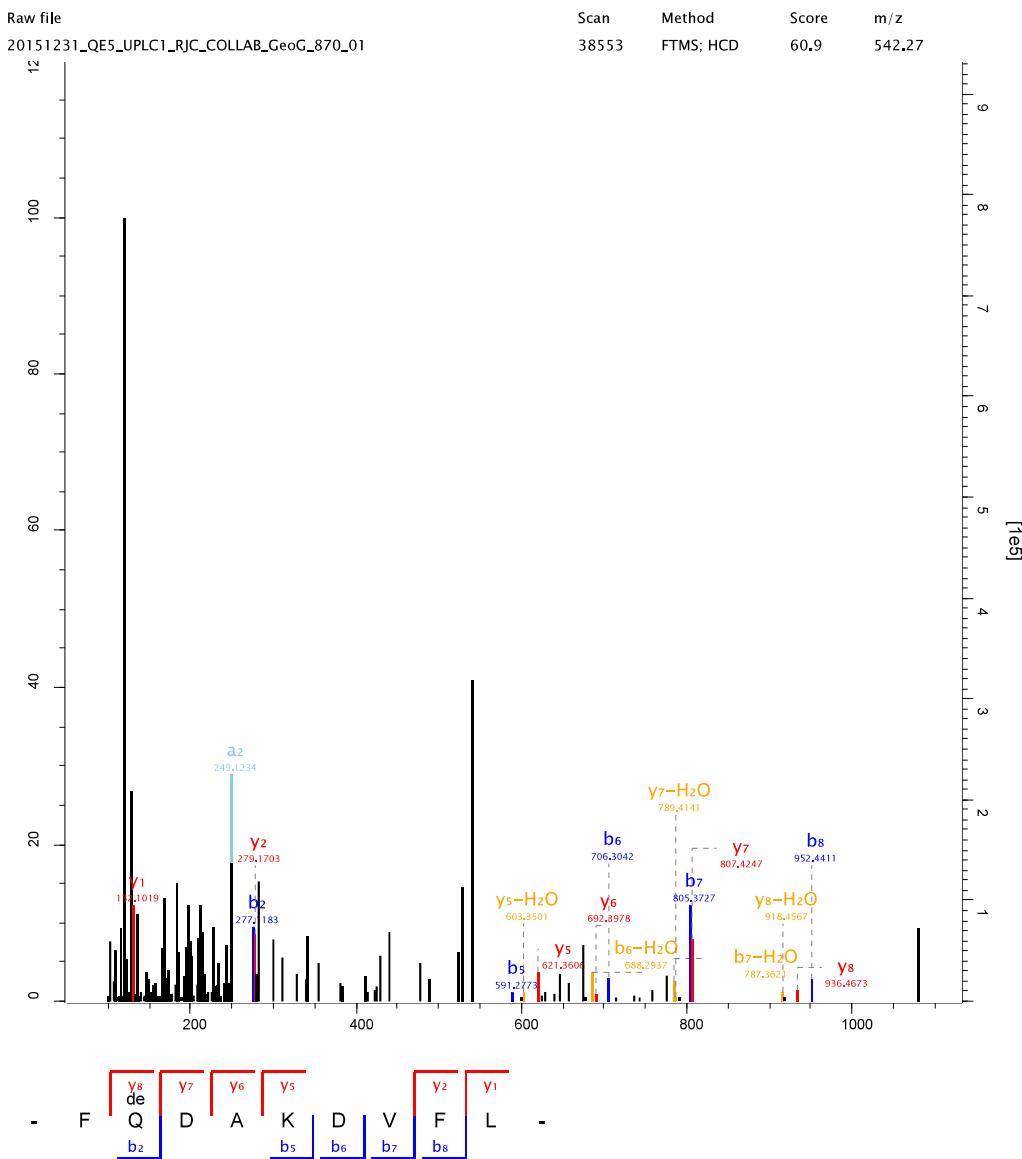


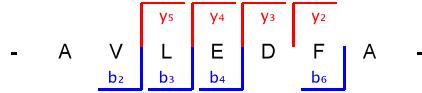
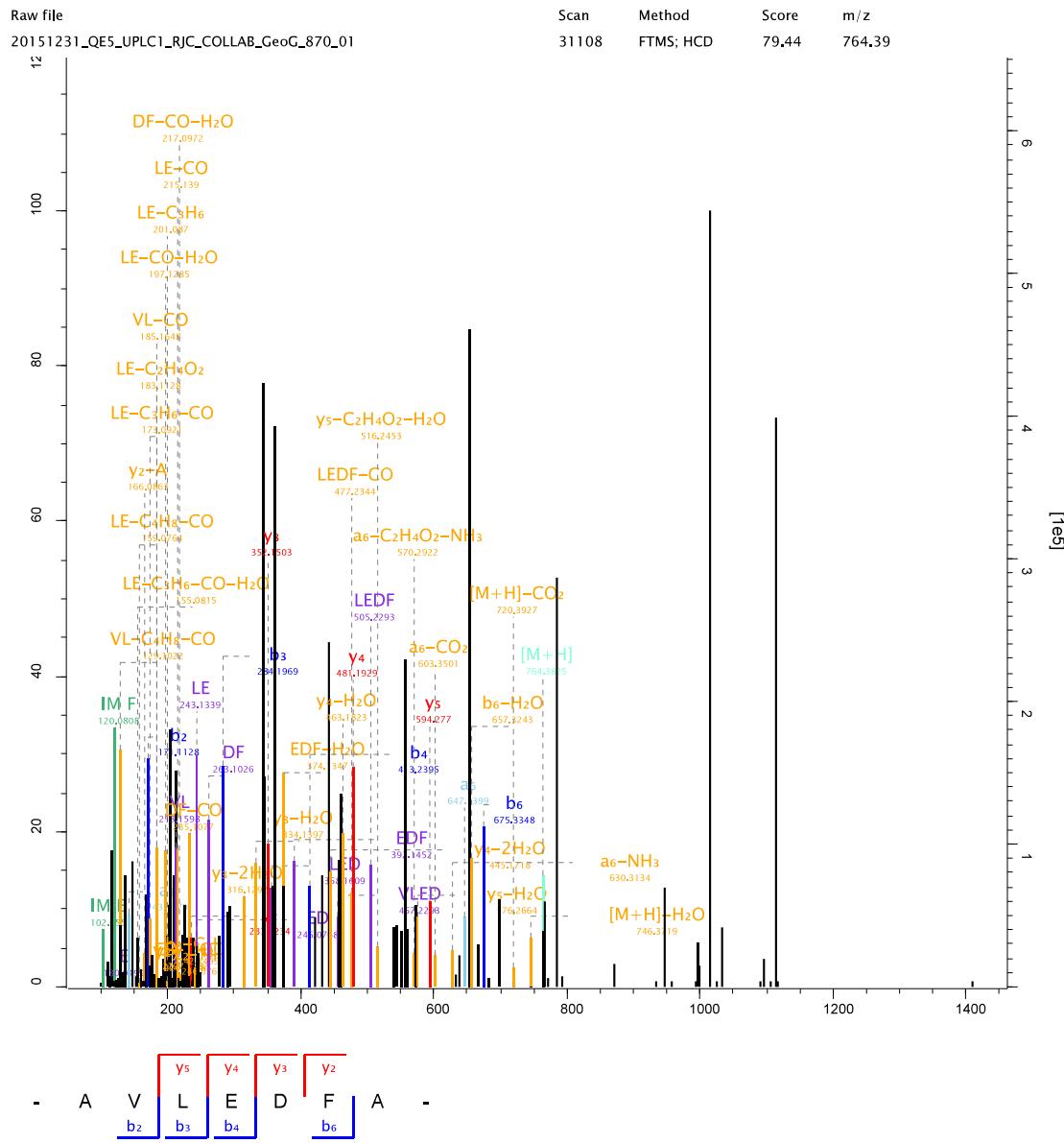


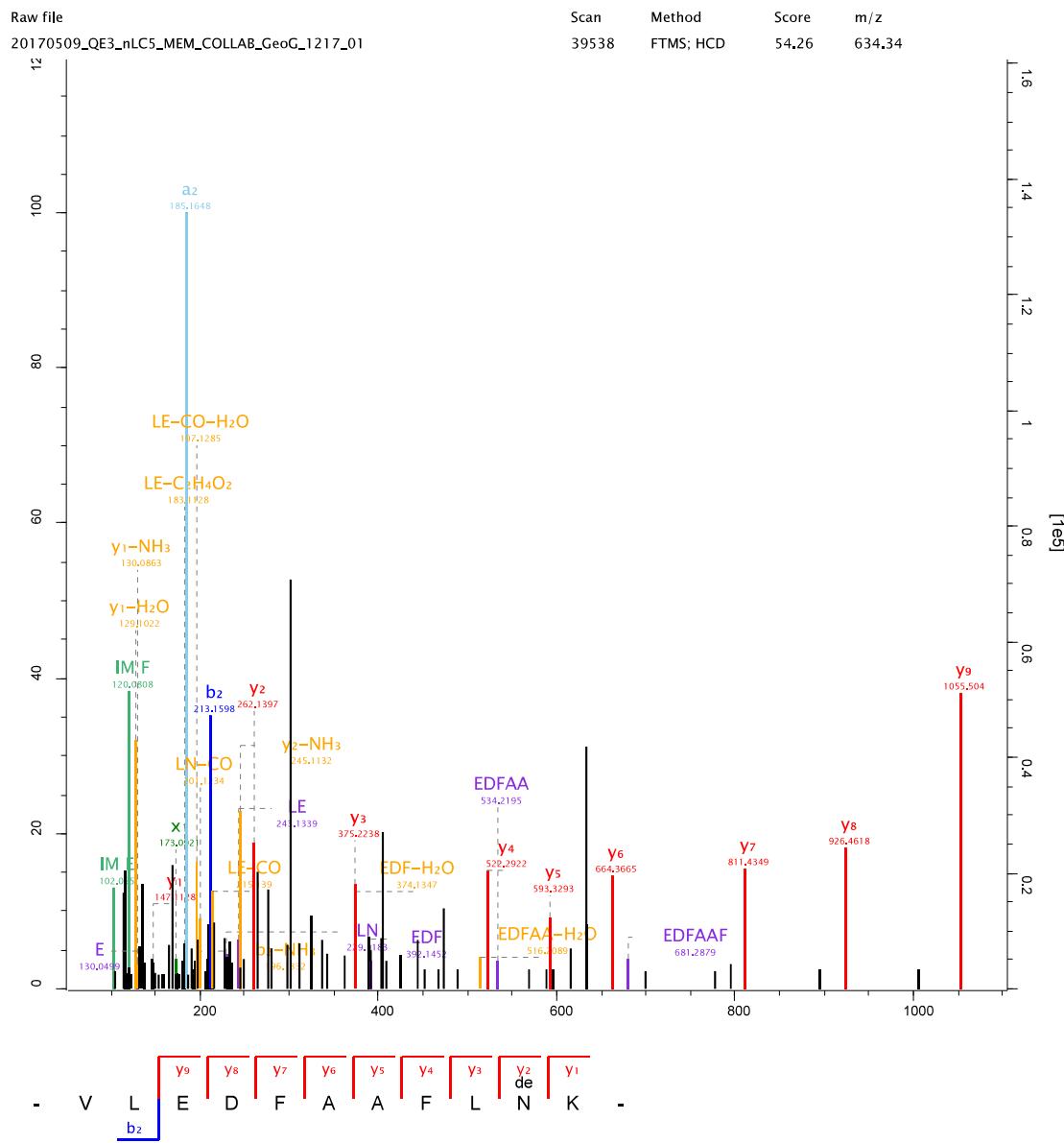
- L K P L V D E P V - N
 b₂ y₈
 b₅ y₃
 b₆ y₂
 b₇ y₁
 b₈ de

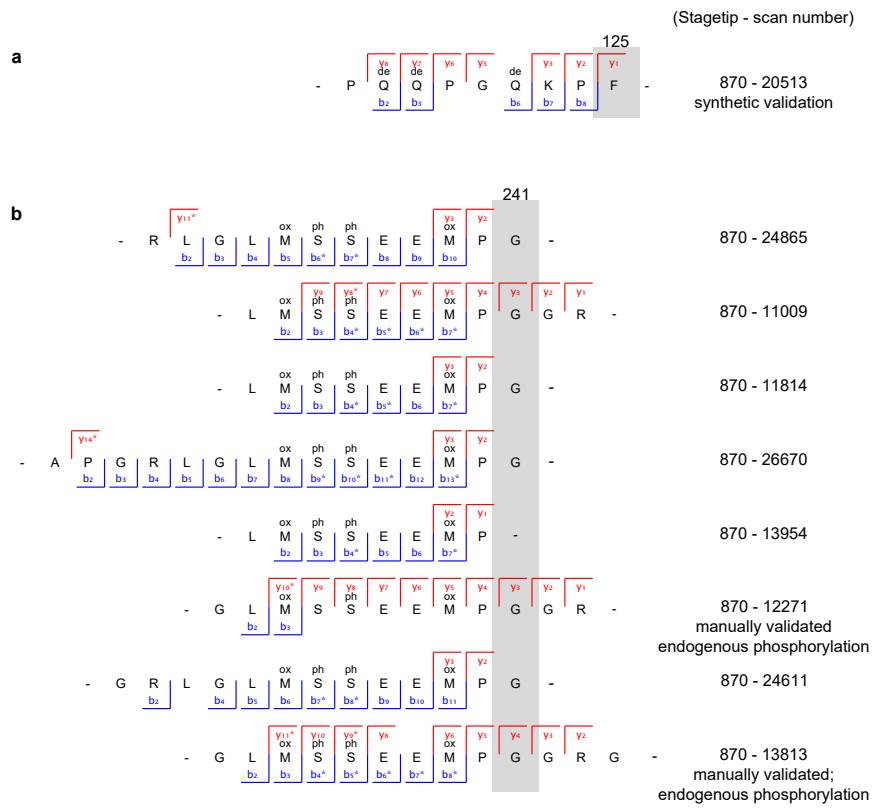


gl	E	y_{12}	y_{11}	y_{10}	y_9	y_8	y_7	y_6	y_5	y_4	y_3	y_2	y_1
-	L	K	P	L	V	D	E	P	V	N	L	V	K
-	b_2	a_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	-	-

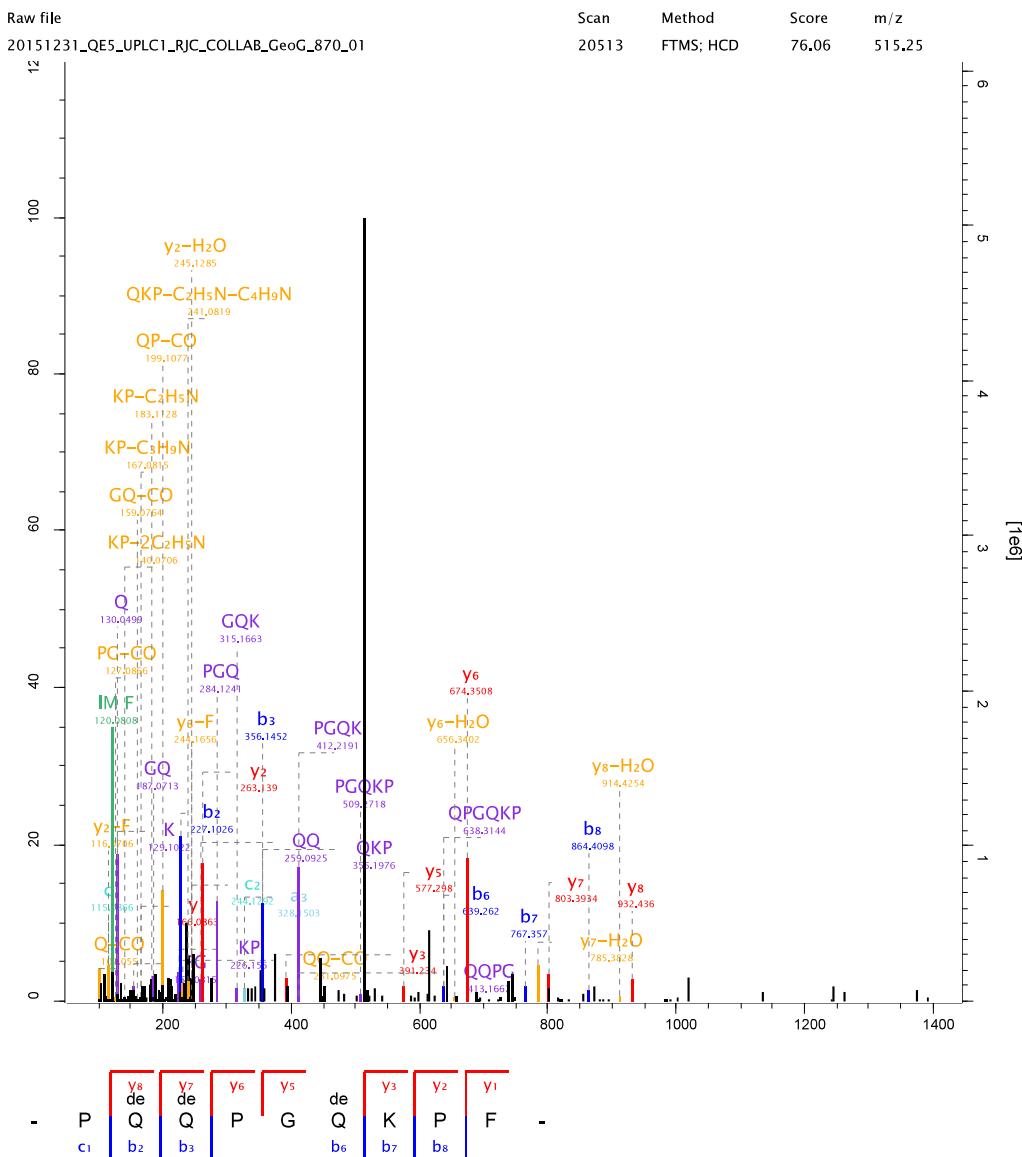


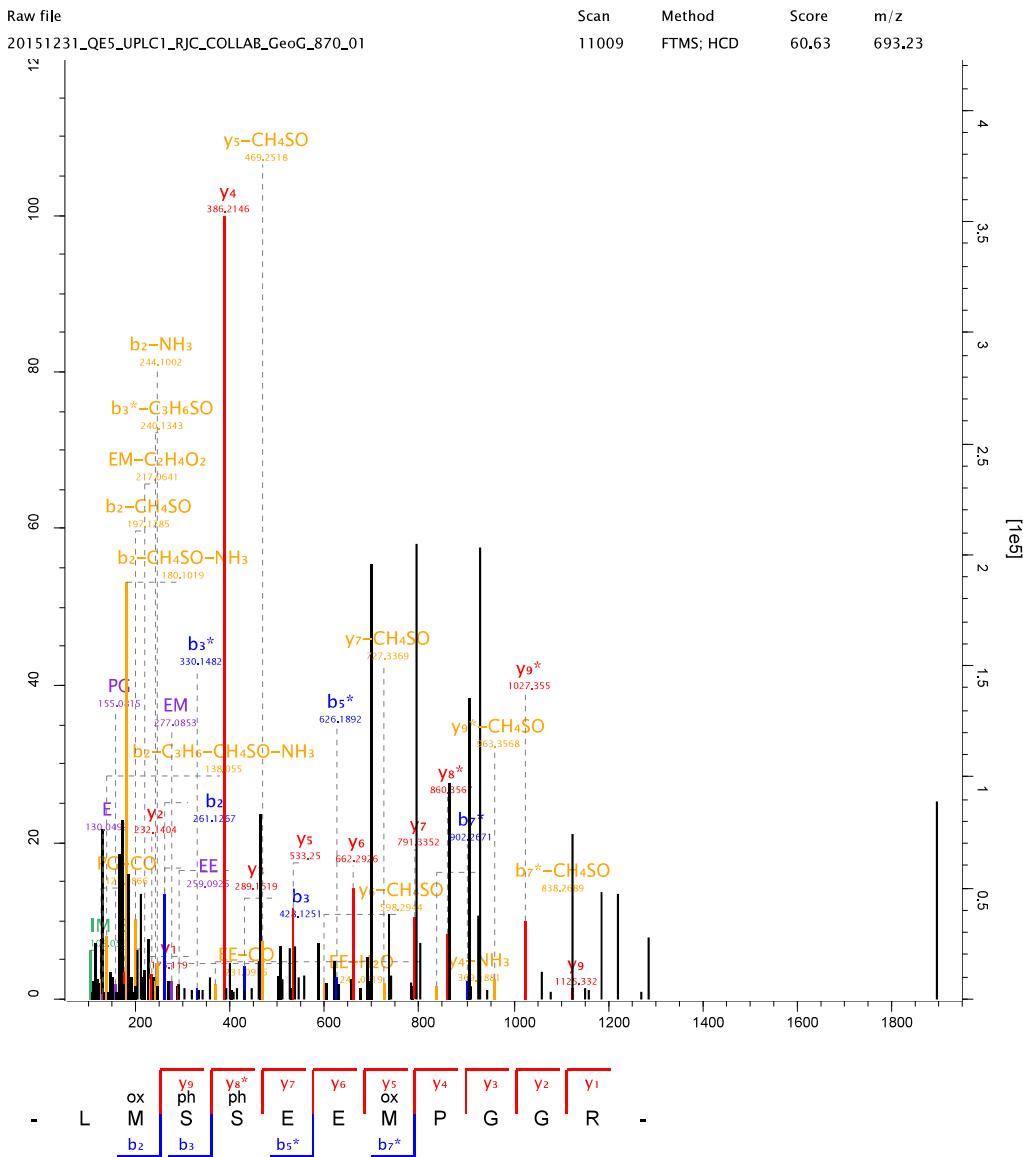


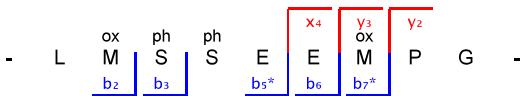
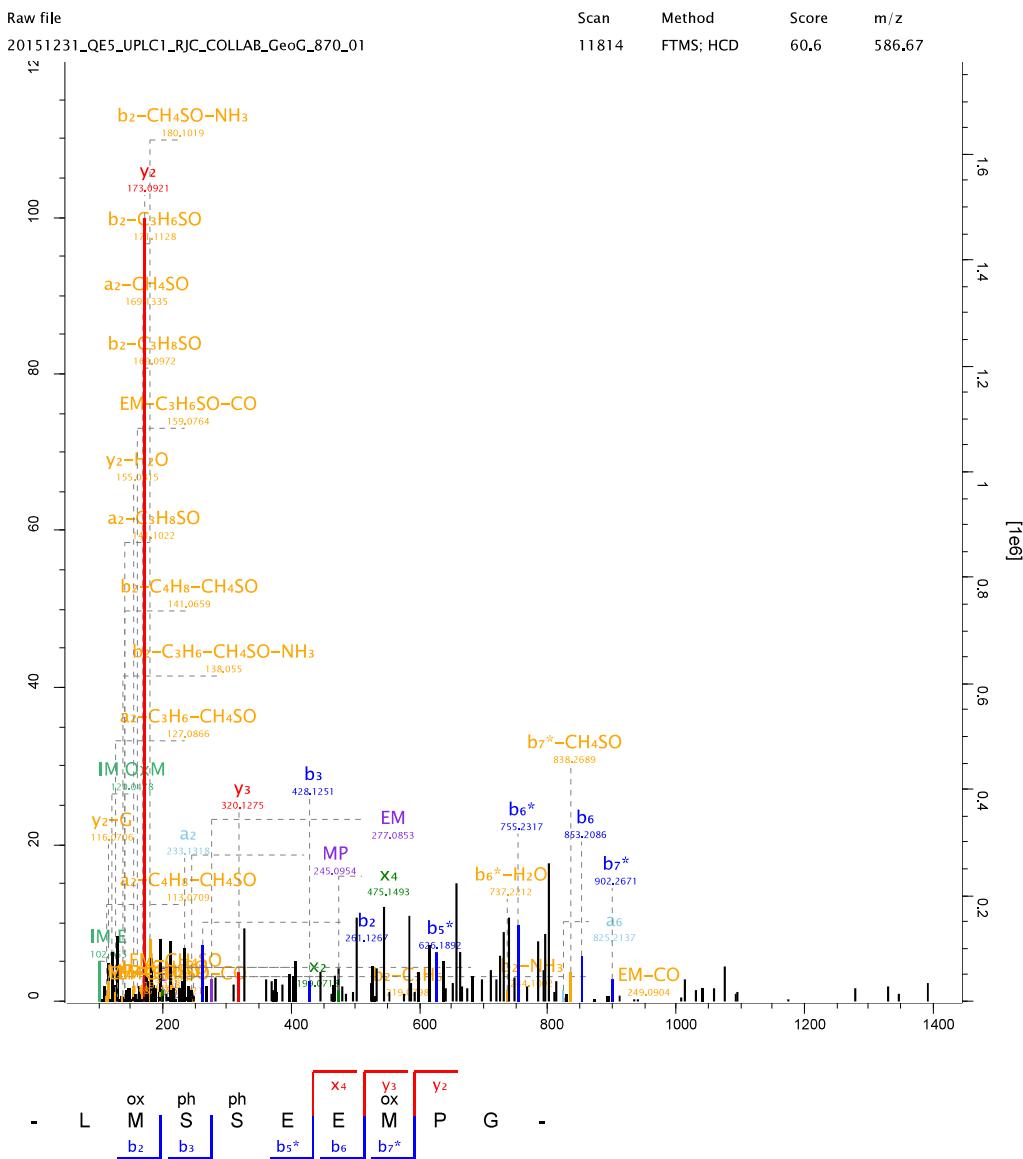


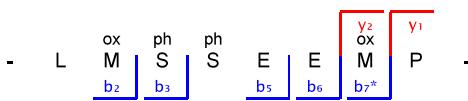
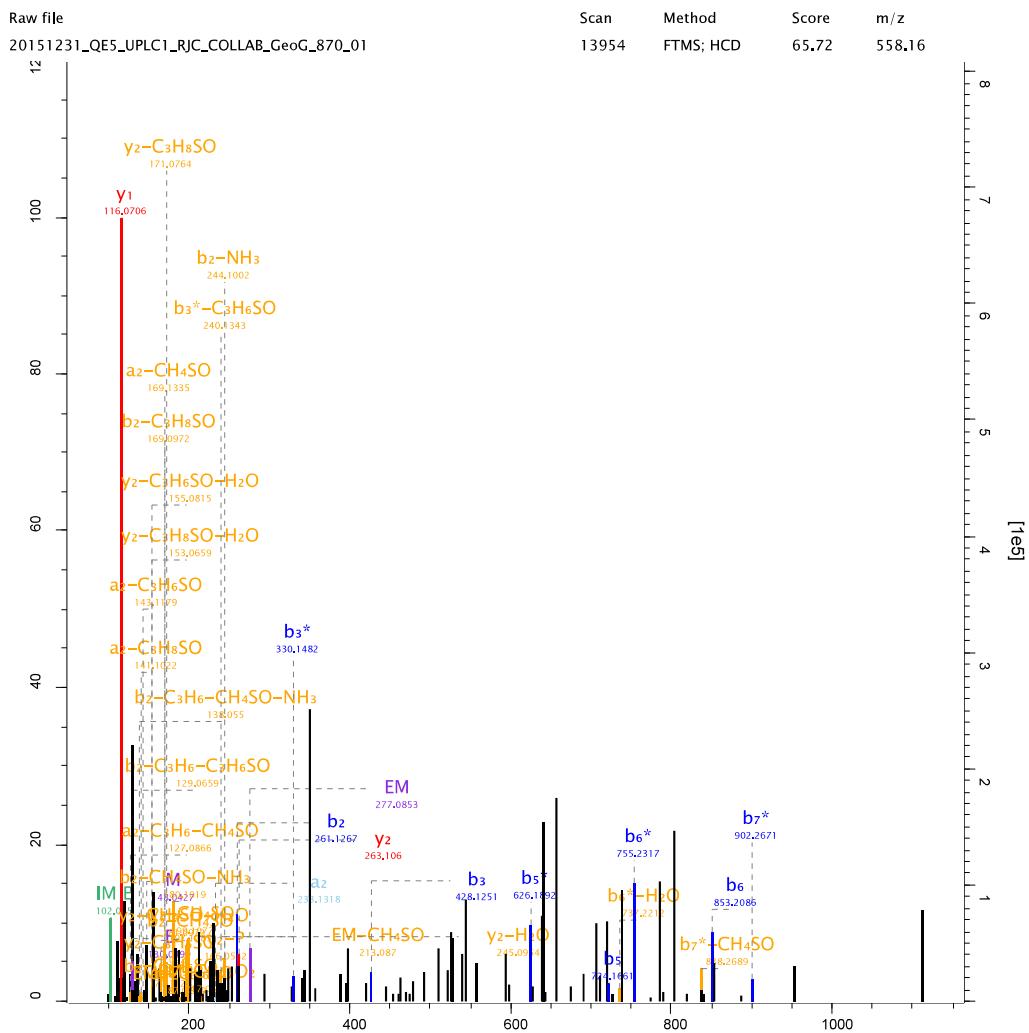


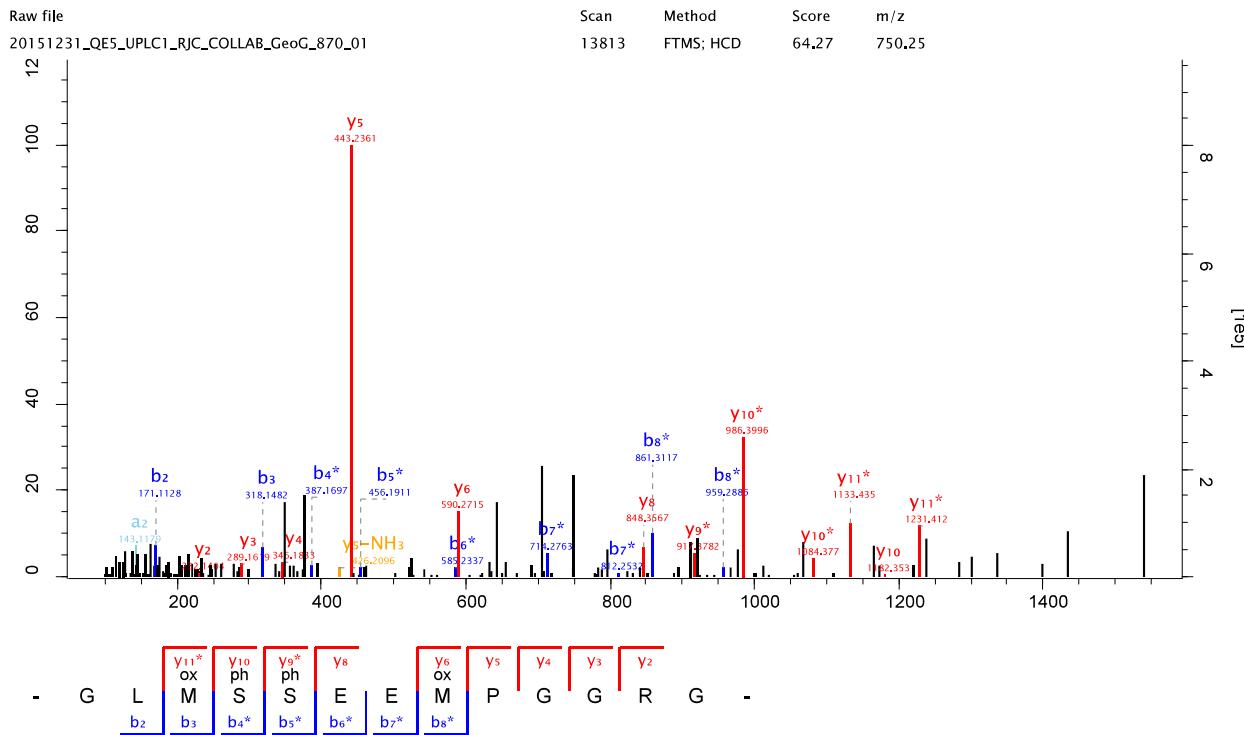
AMB. Peptide matches to phylogenetically informative positions 125 (a), and 241 (b). Coordinates in reference to XP_004419200.



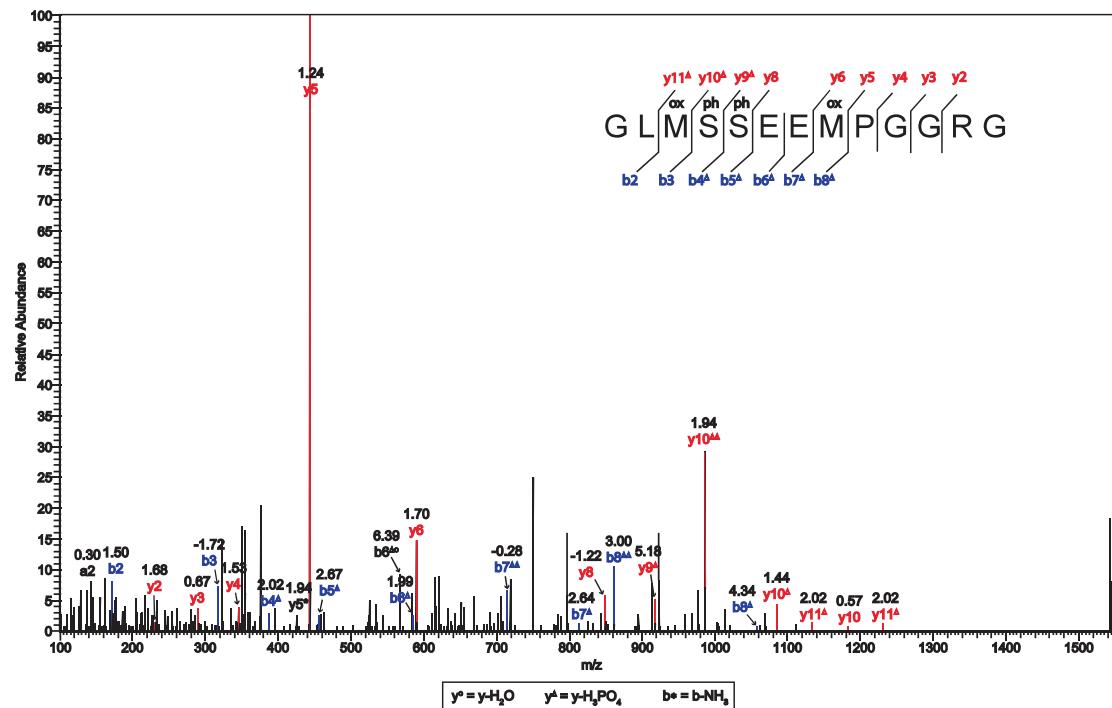


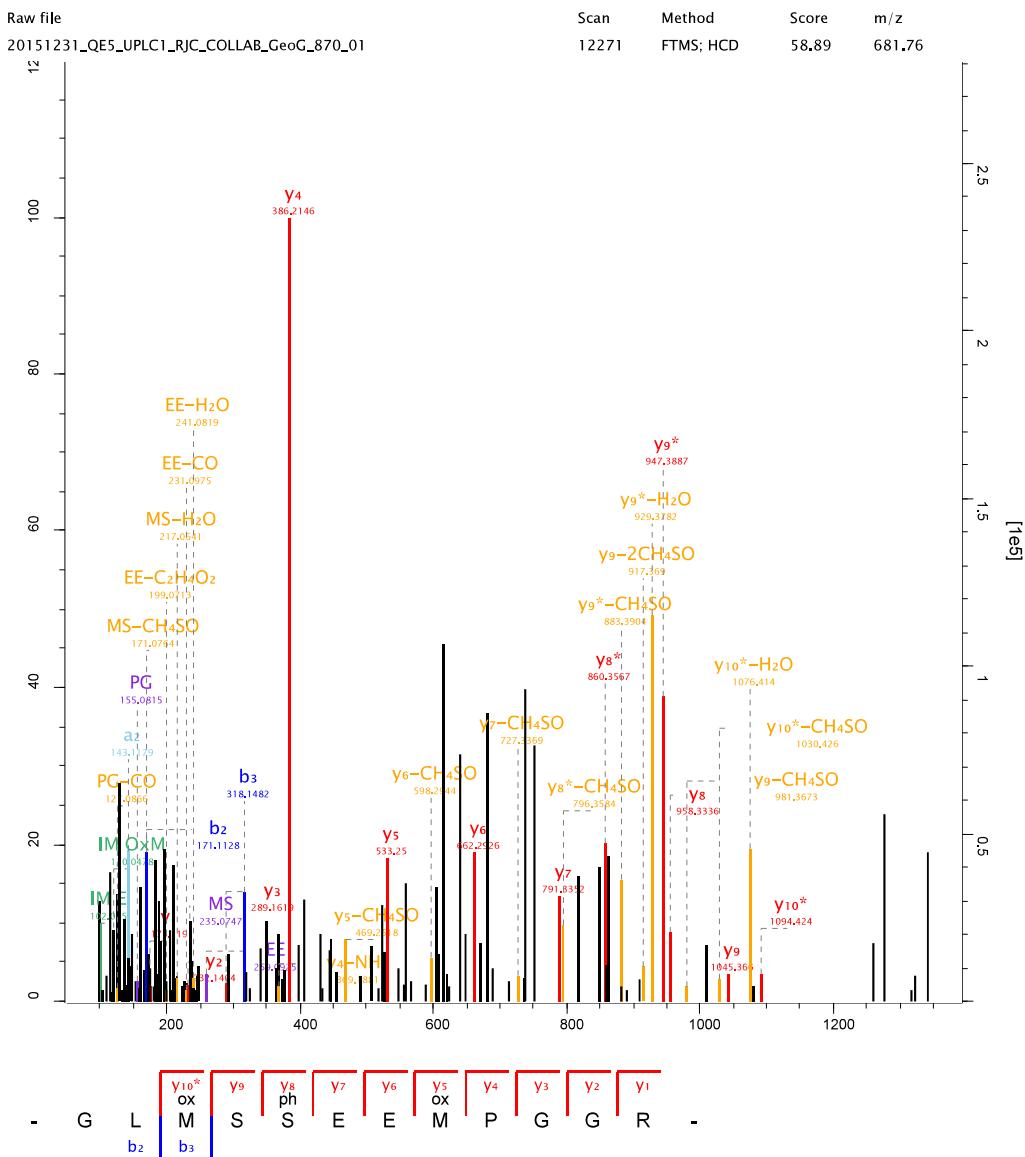




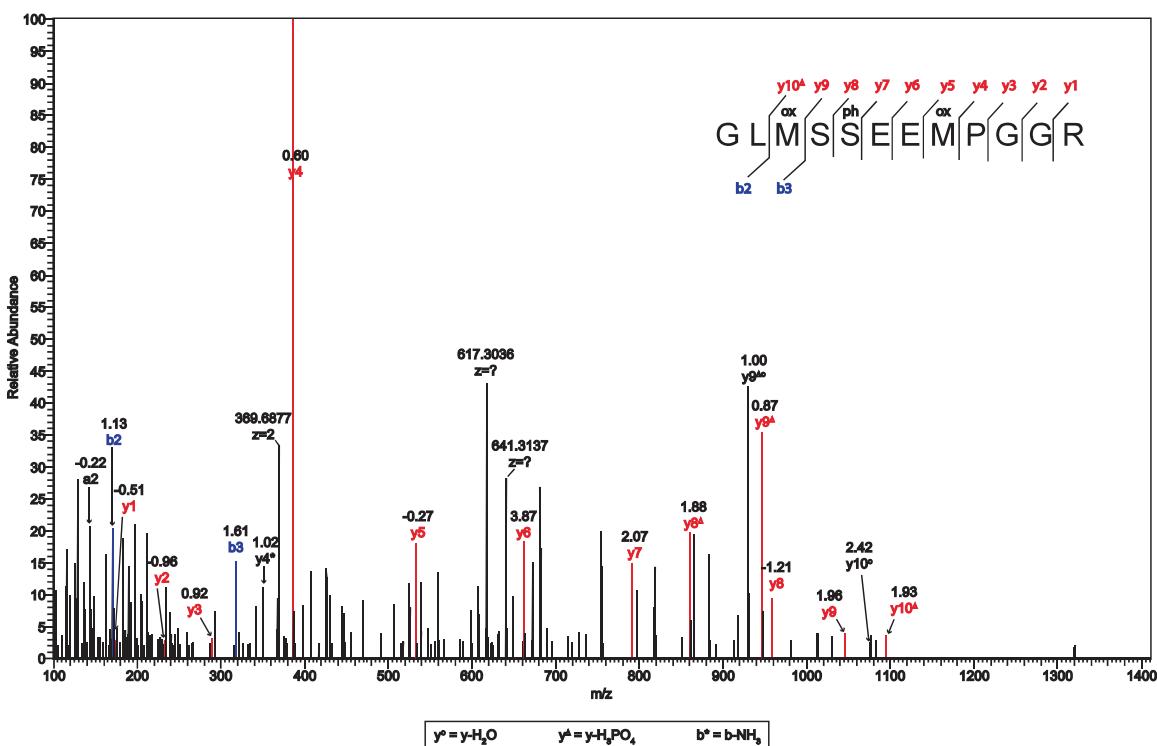


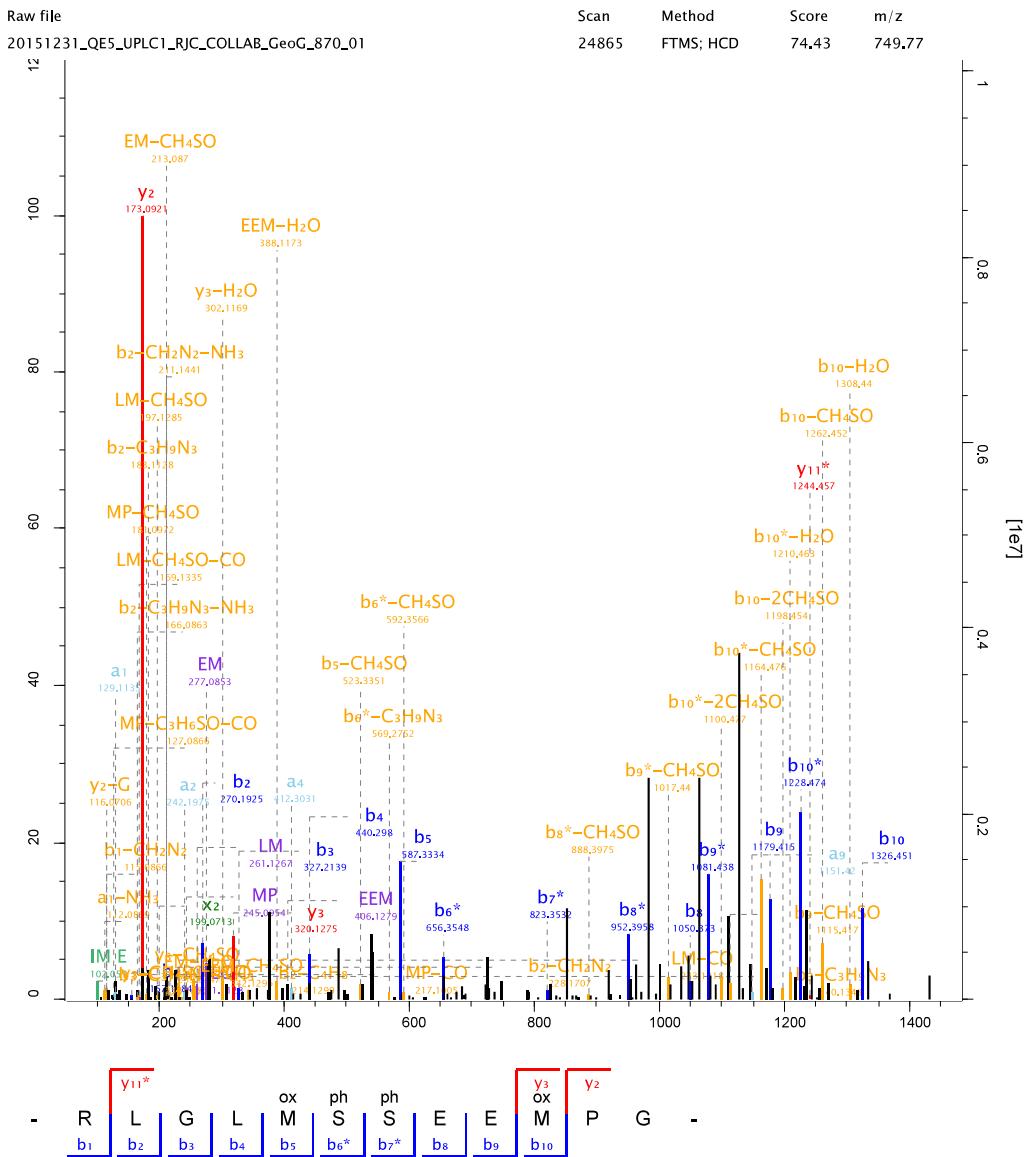
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01 #13813 RT:52.32 AV:1 NL:1.87E5
T: FTMS +p NSId Full ms2 750.25@hcd28.00 [f00.00-1550.00]

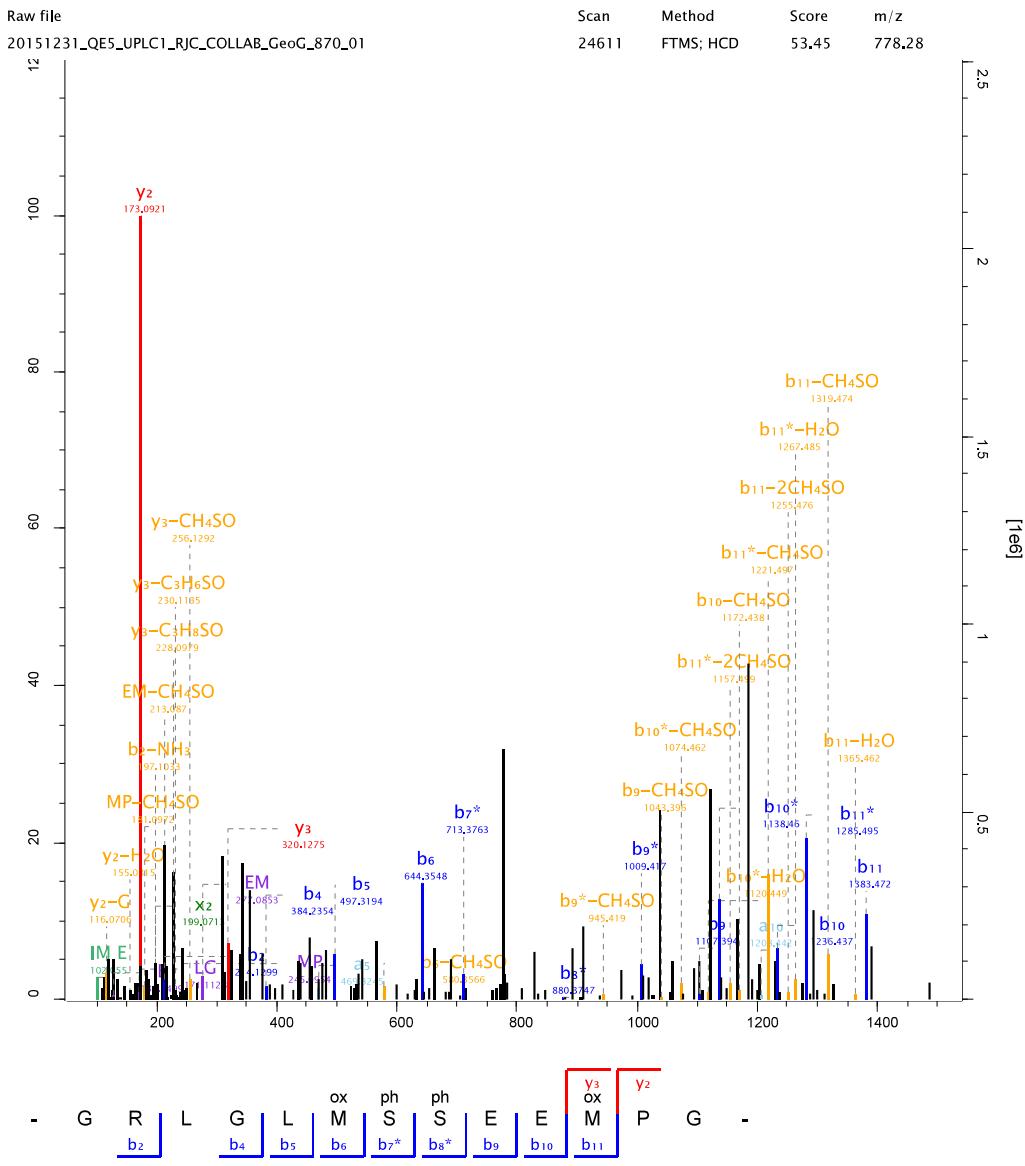


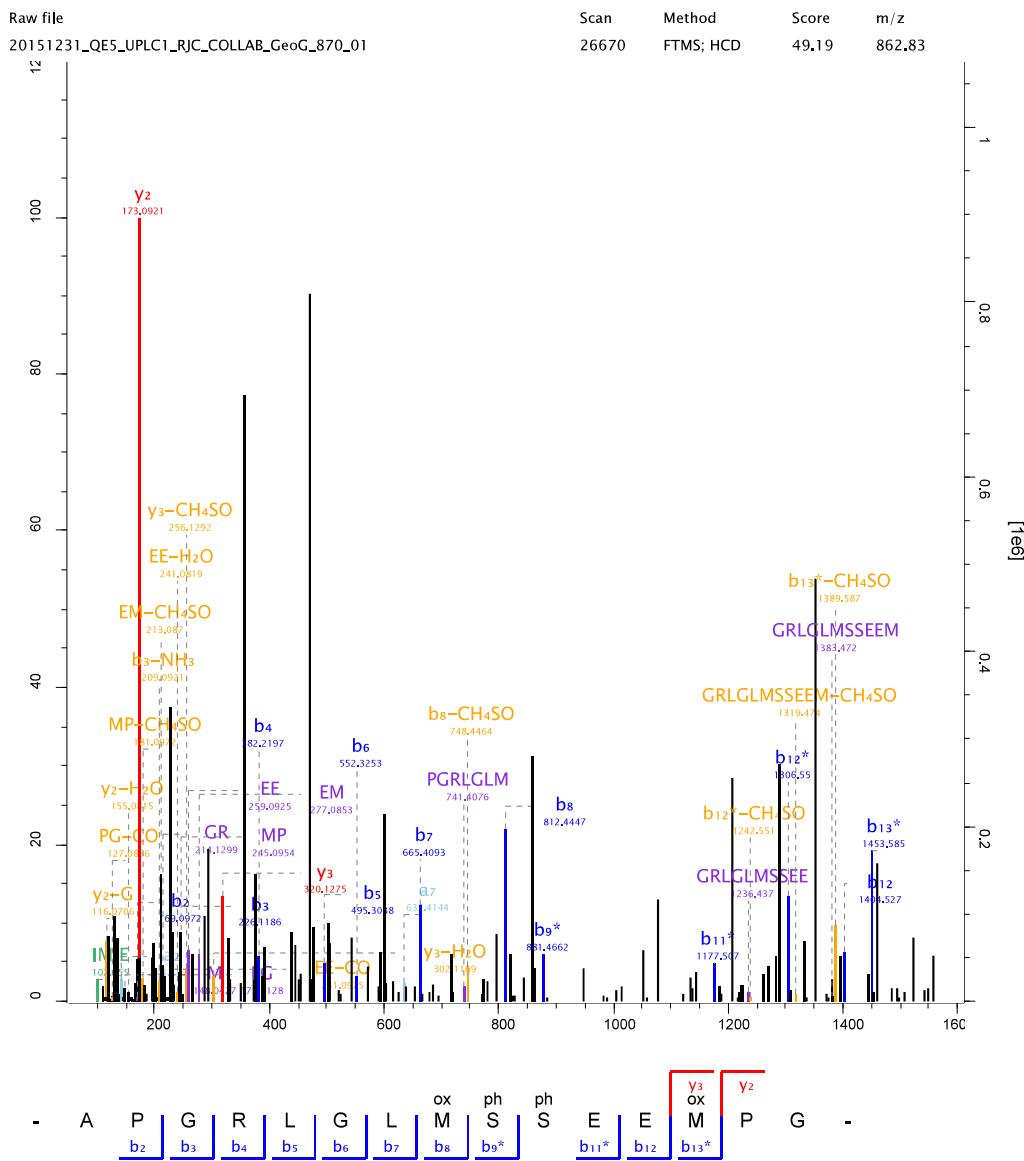


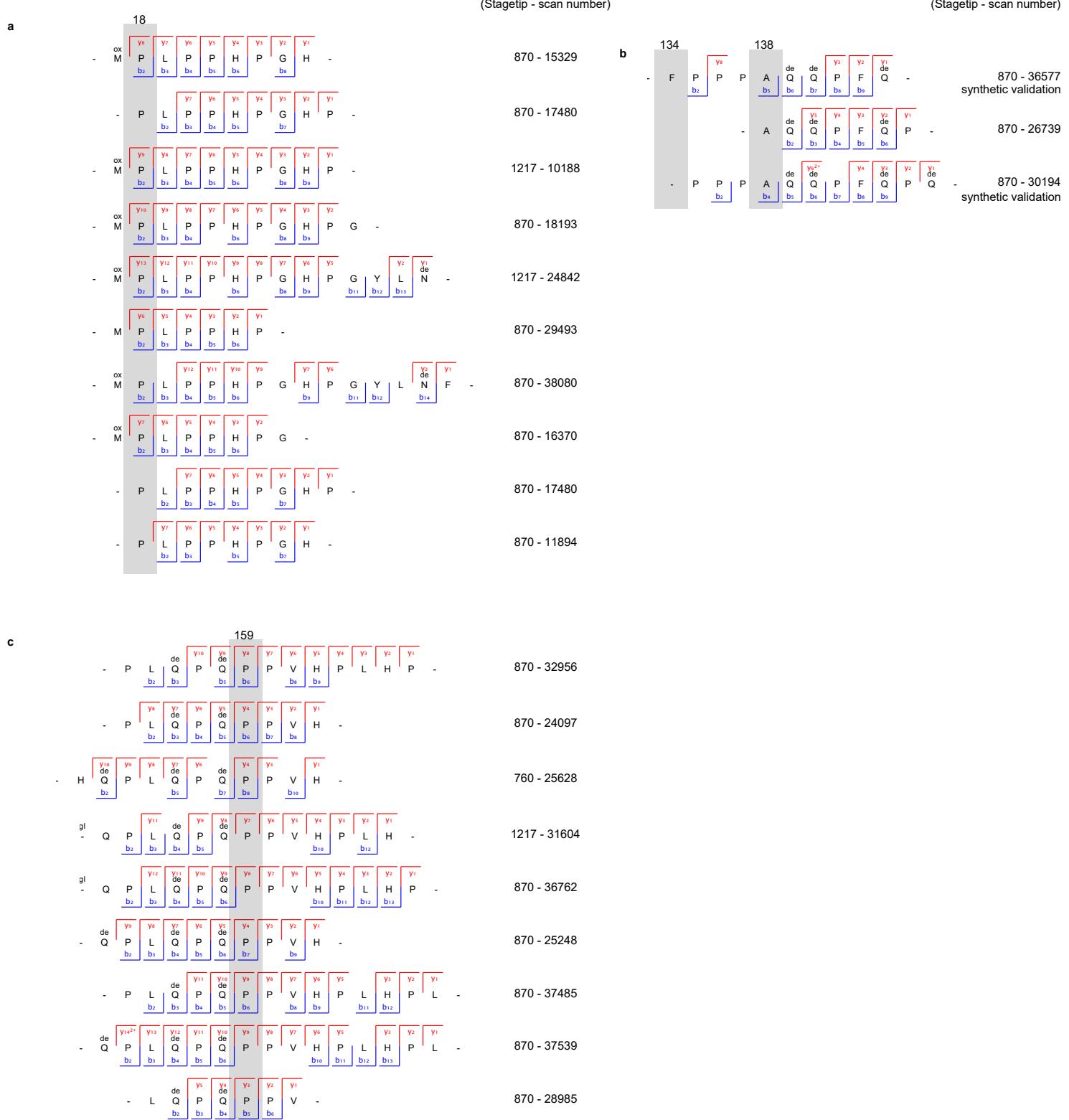
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01 #12271 RT: 48.29 AV: 1 NL: 5.70E4
T: FTMS +p NSId Full ms2 681.76@hcd28.00 [100.00-1410.00]



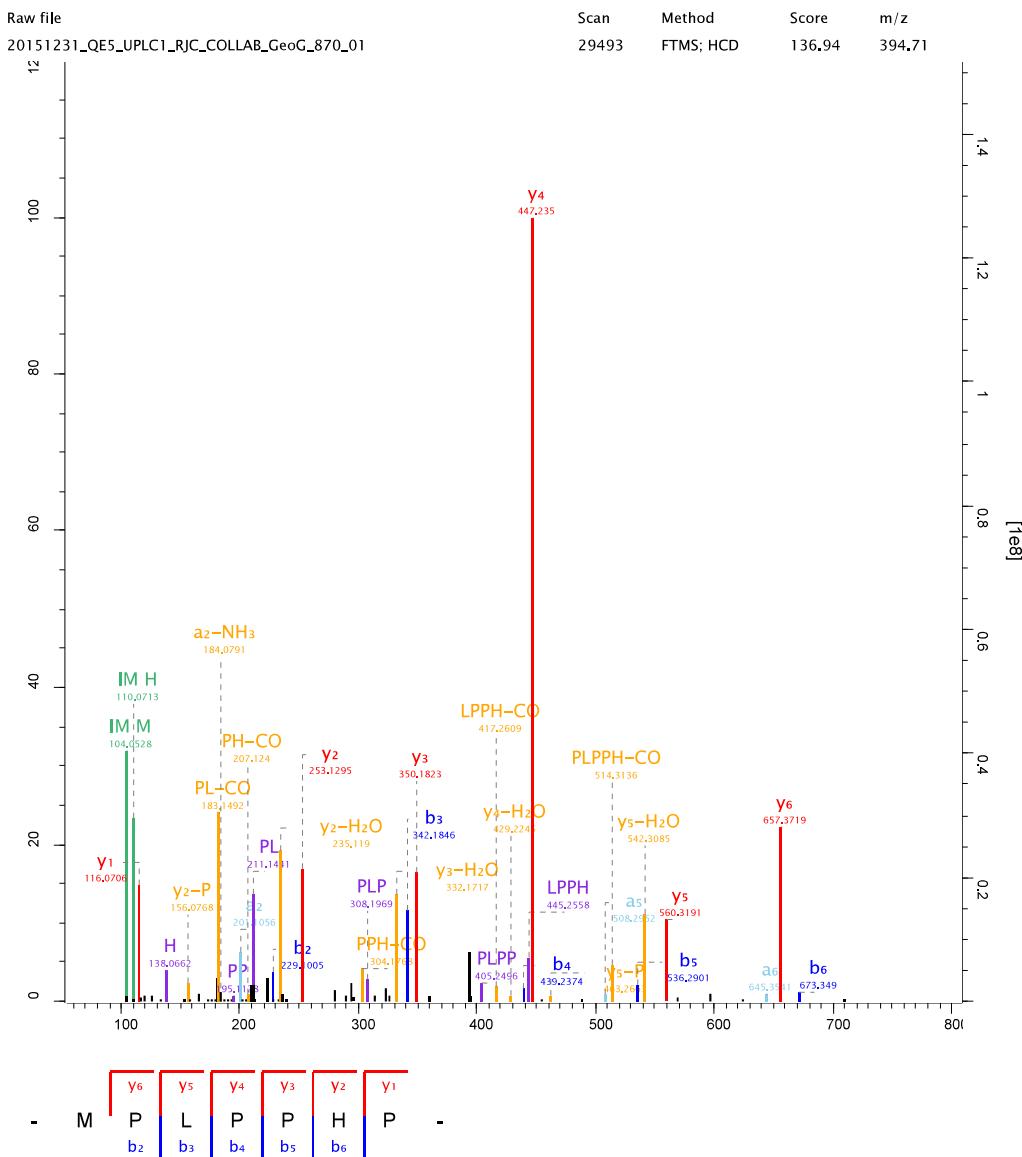


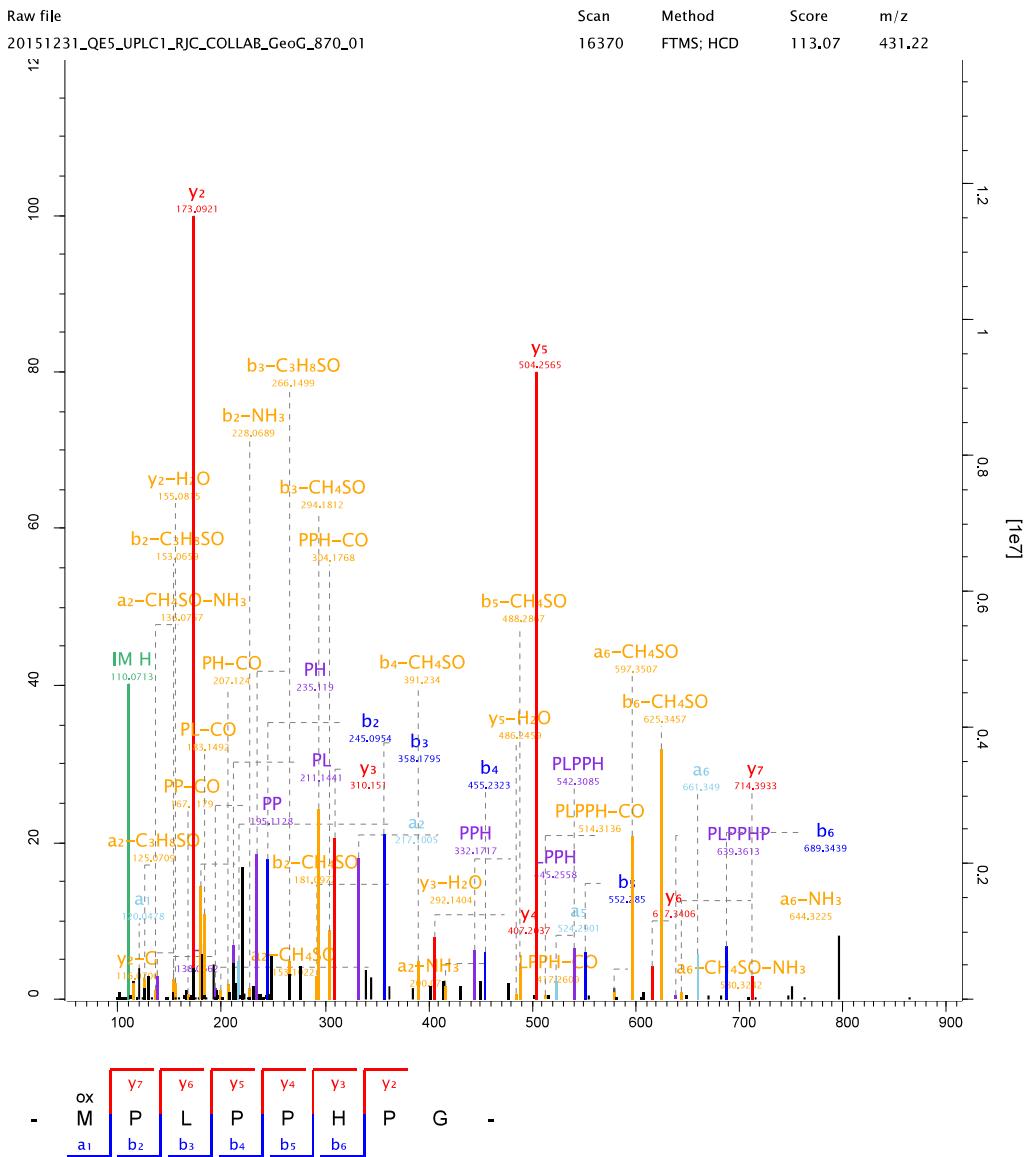


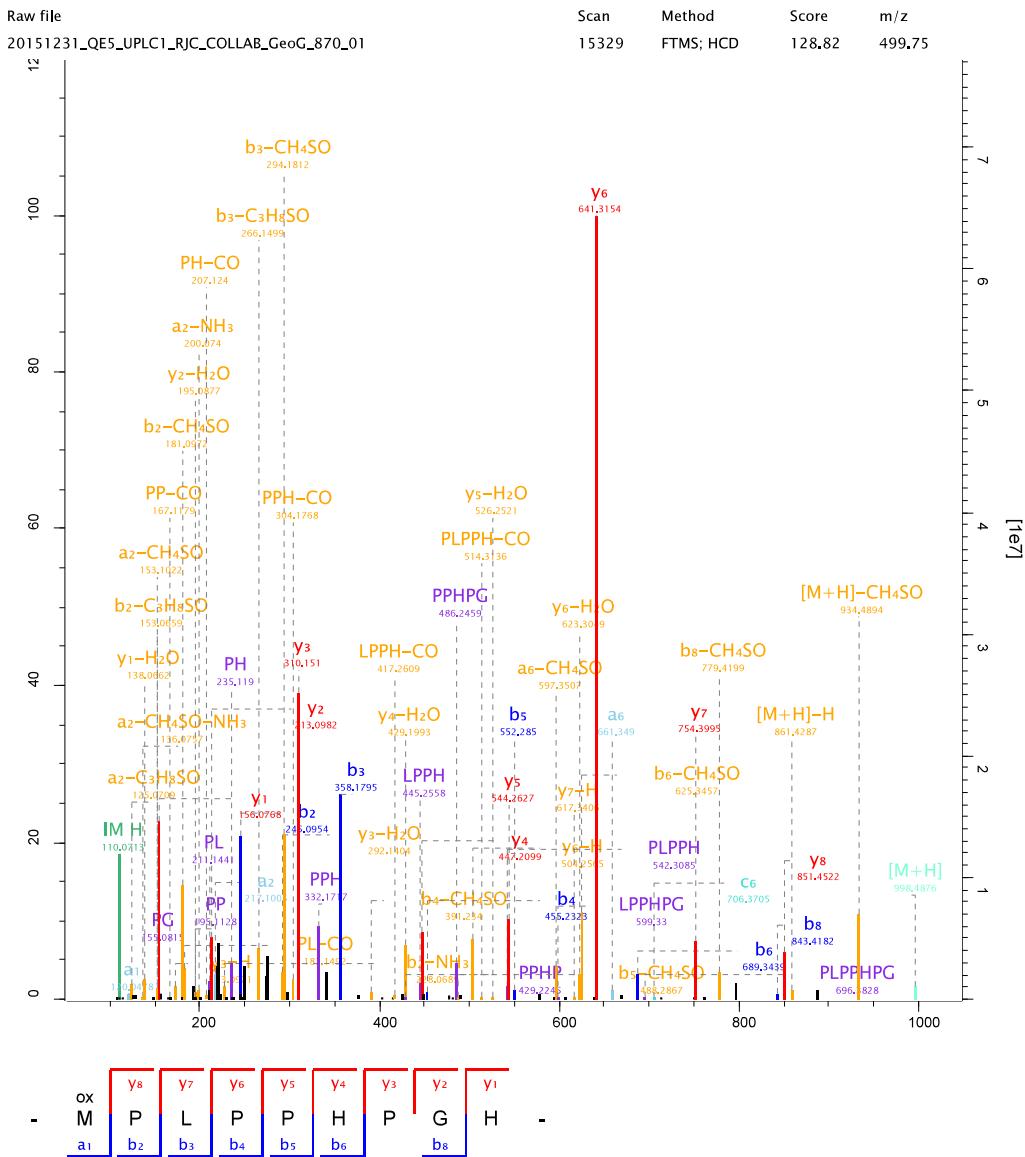


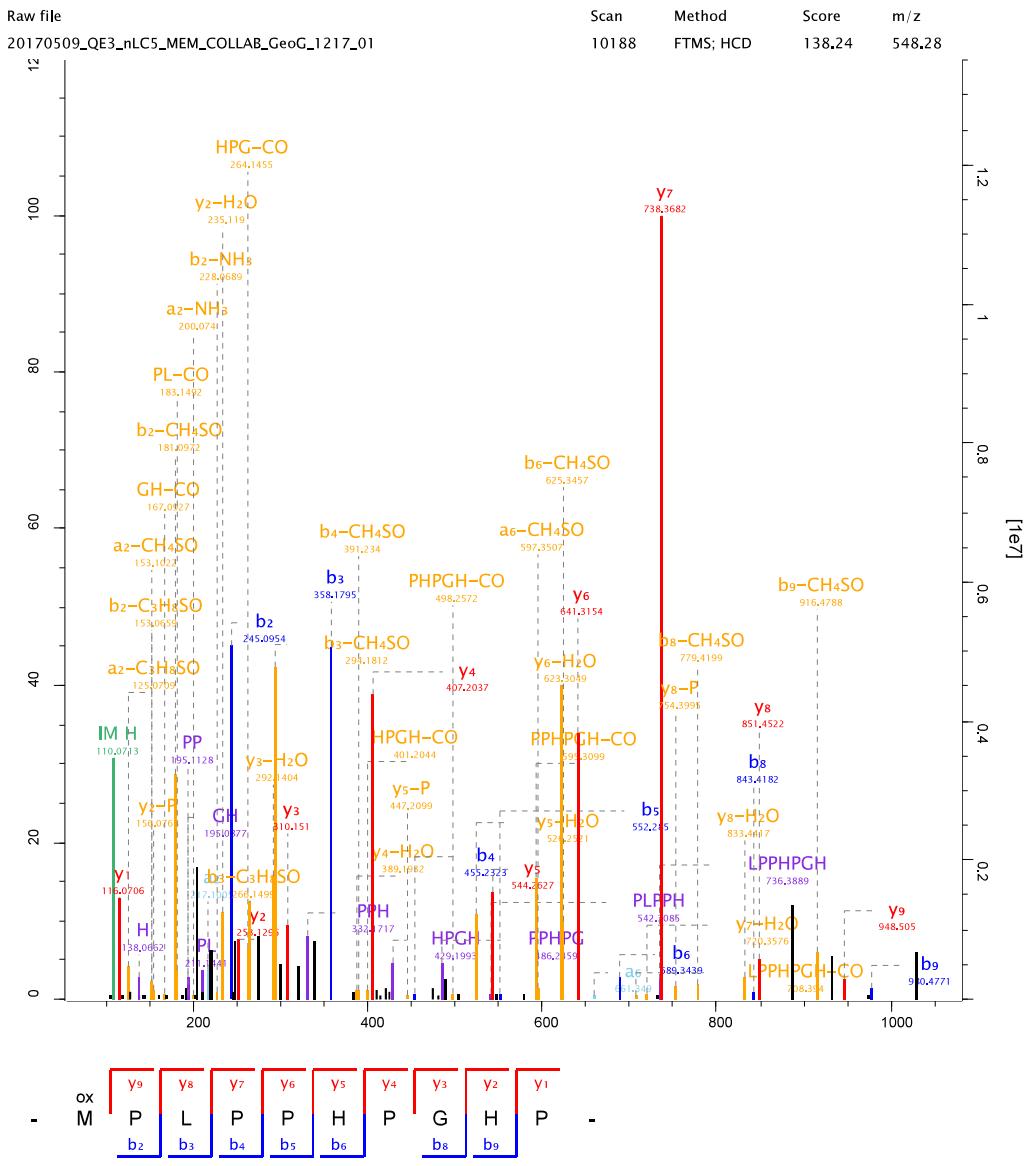


AMELX. Peptide matches to phylogenetically informative positions 18 (a), 134, 138 (b), and 159 (c). Coordinates in reference to XP_004435161.



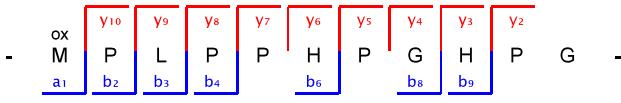
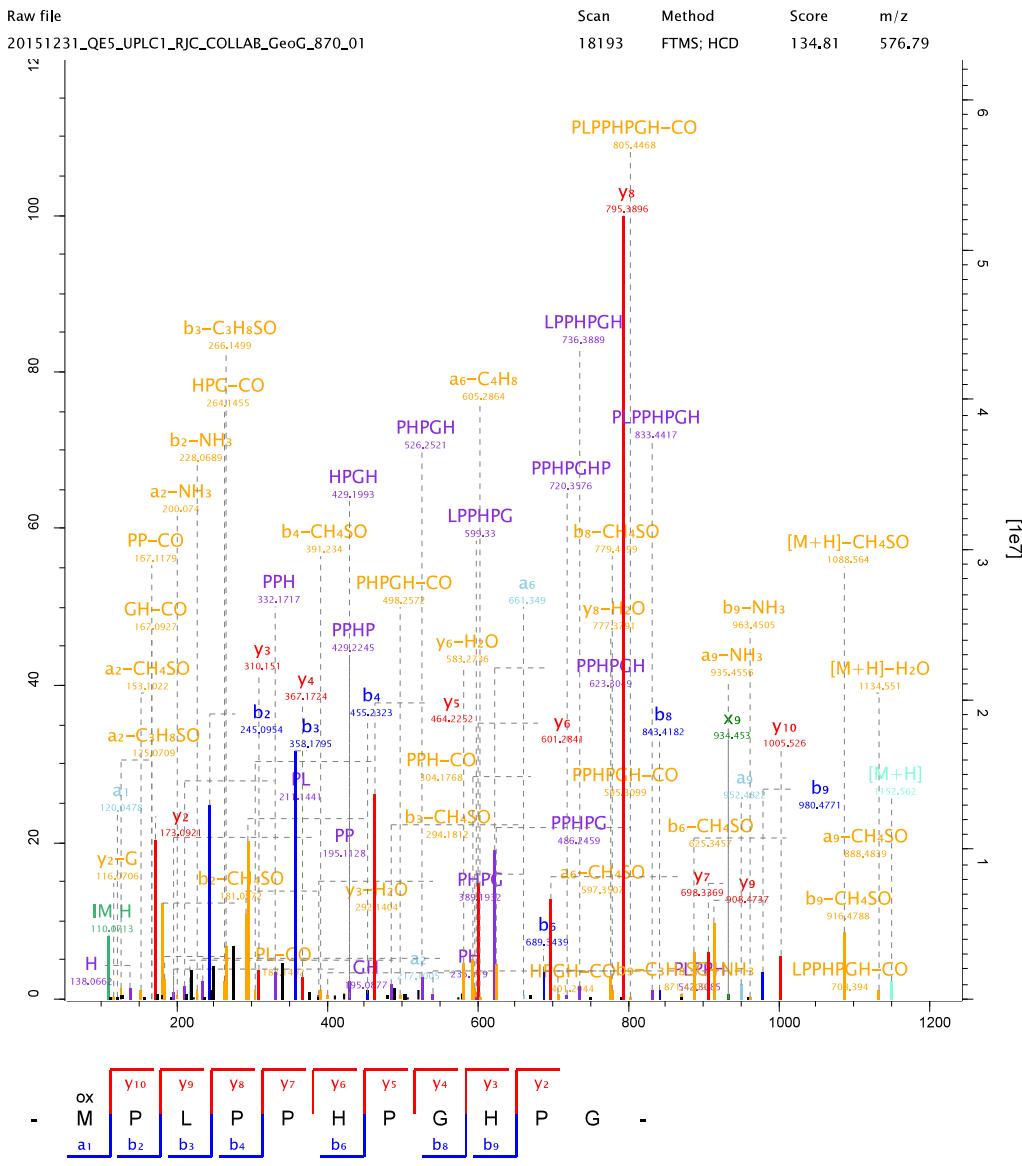


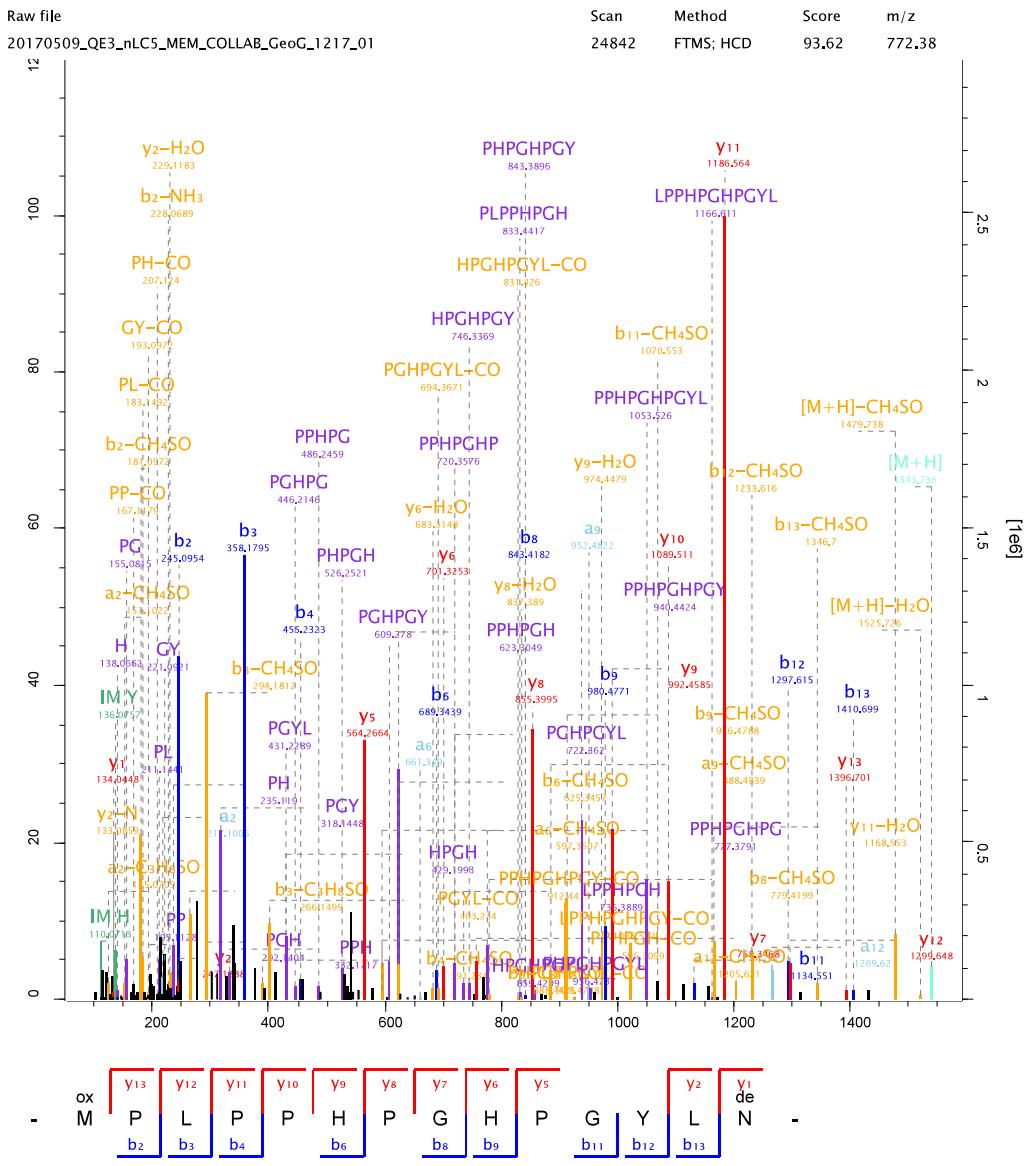




ox M - P L P H P G H P

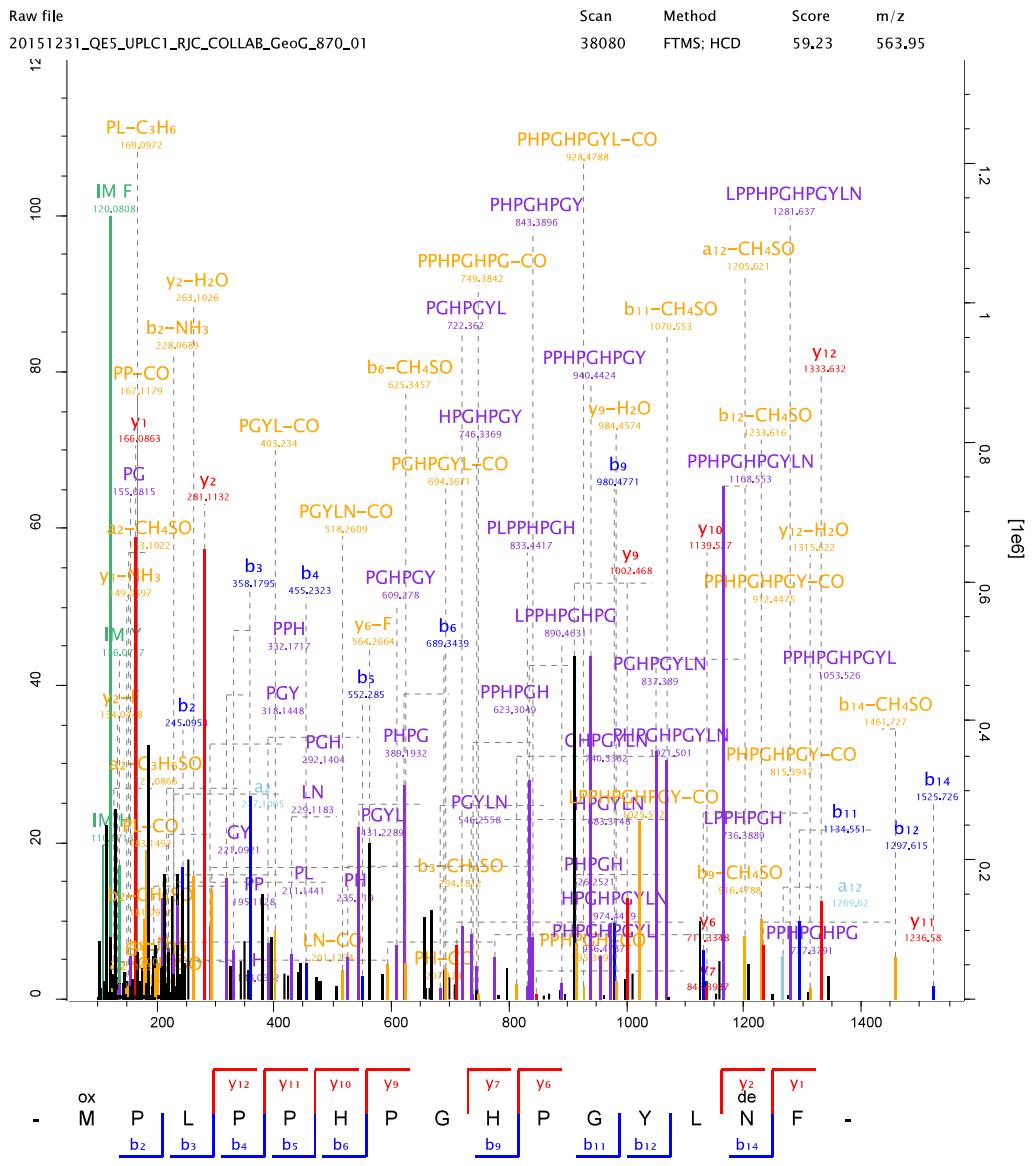
b₂ b₃ b₄ b₅ b₆ b₇ b₈ b₉



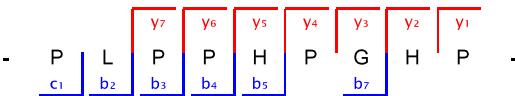
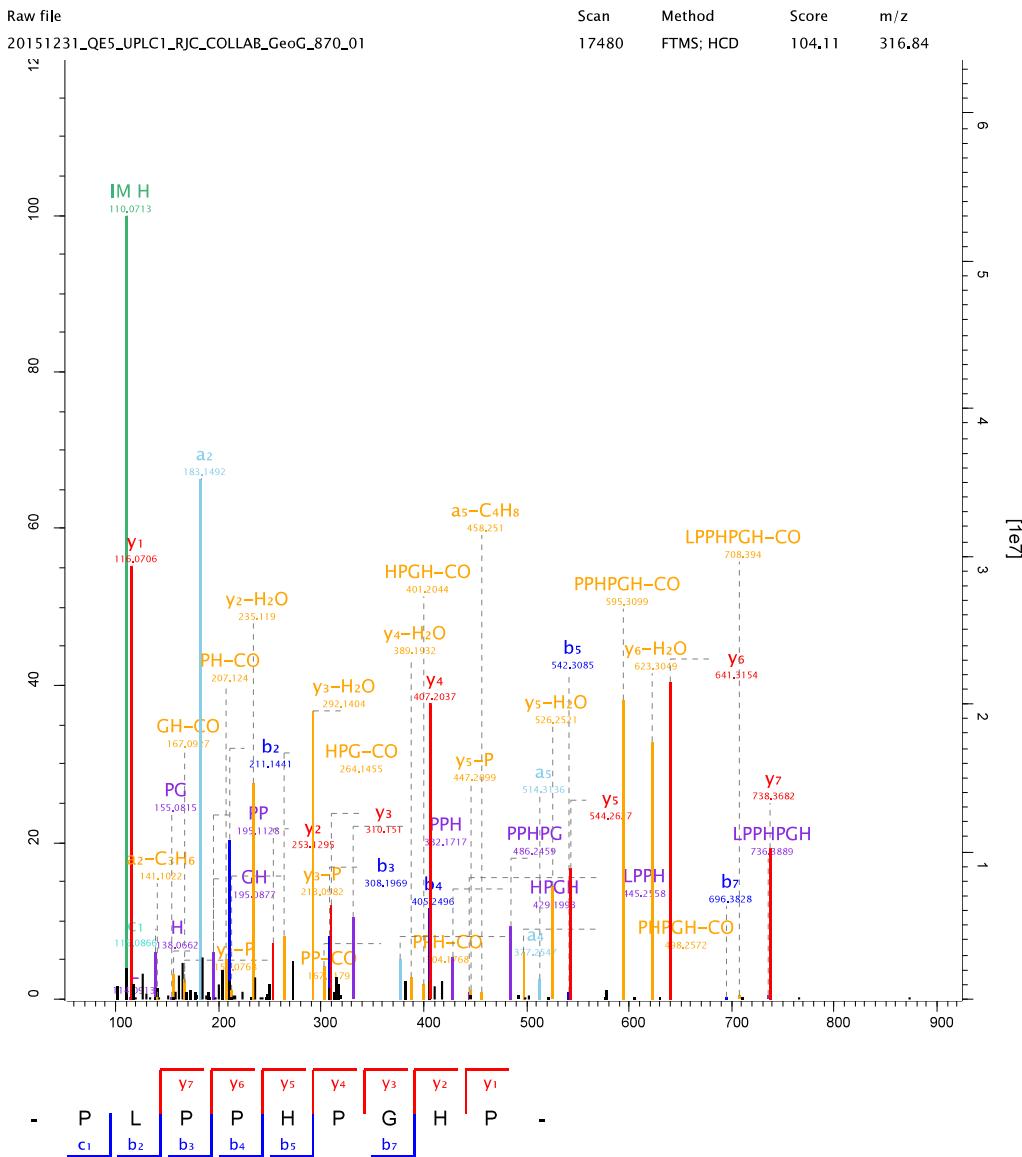


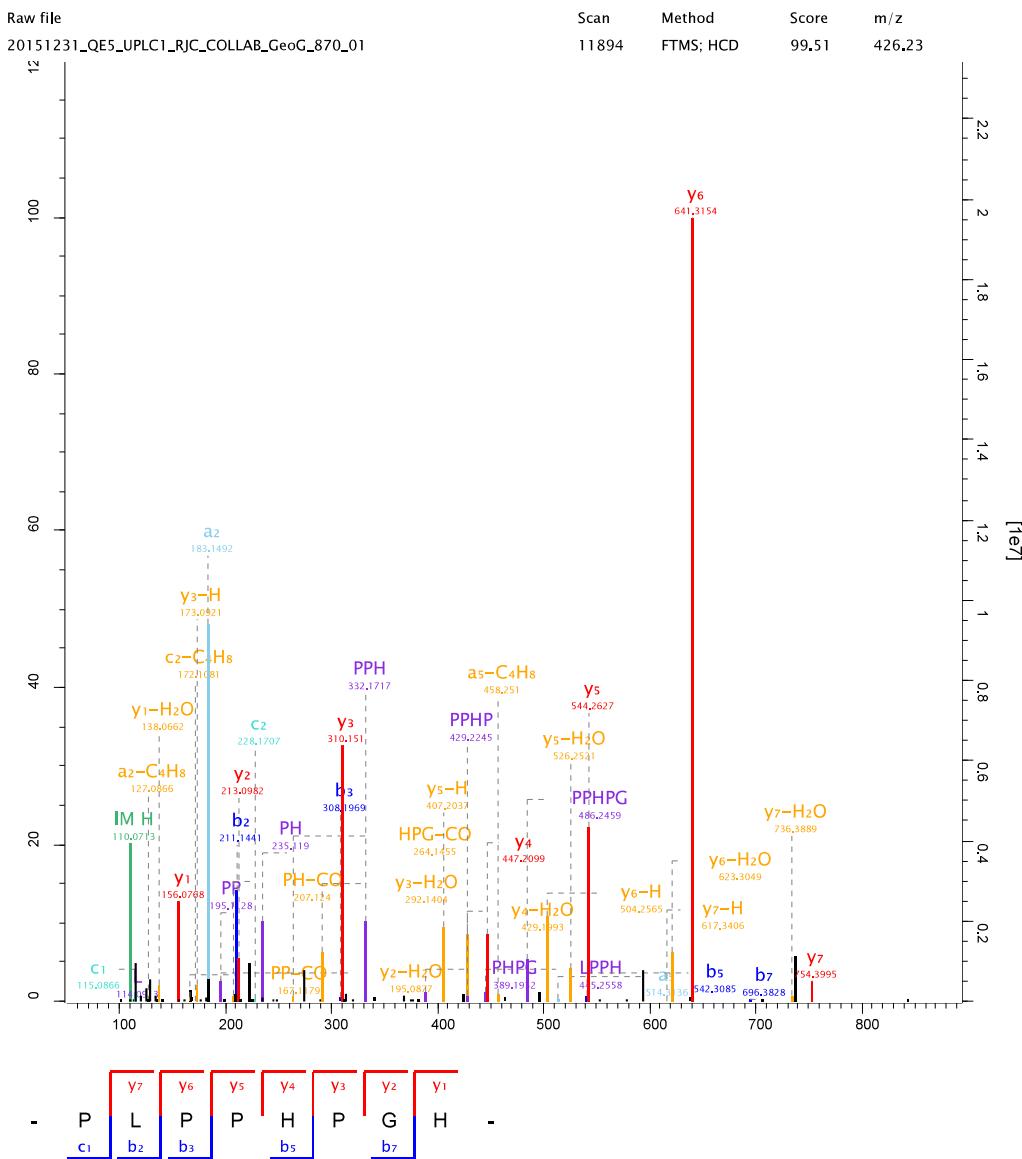
ox M [y₁₃] [y₁₂] [y₁₁] [y₁₀] [y₉] [y₈] [y₇] [y₆] [y₅] P G H Y L N de

b₂ b₃ b₄ b₅ b₆ b₇ b₈ b₉ b₁₁ b₁₂ b₁₃



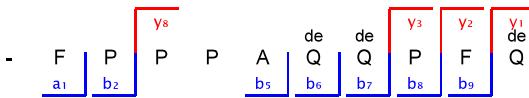
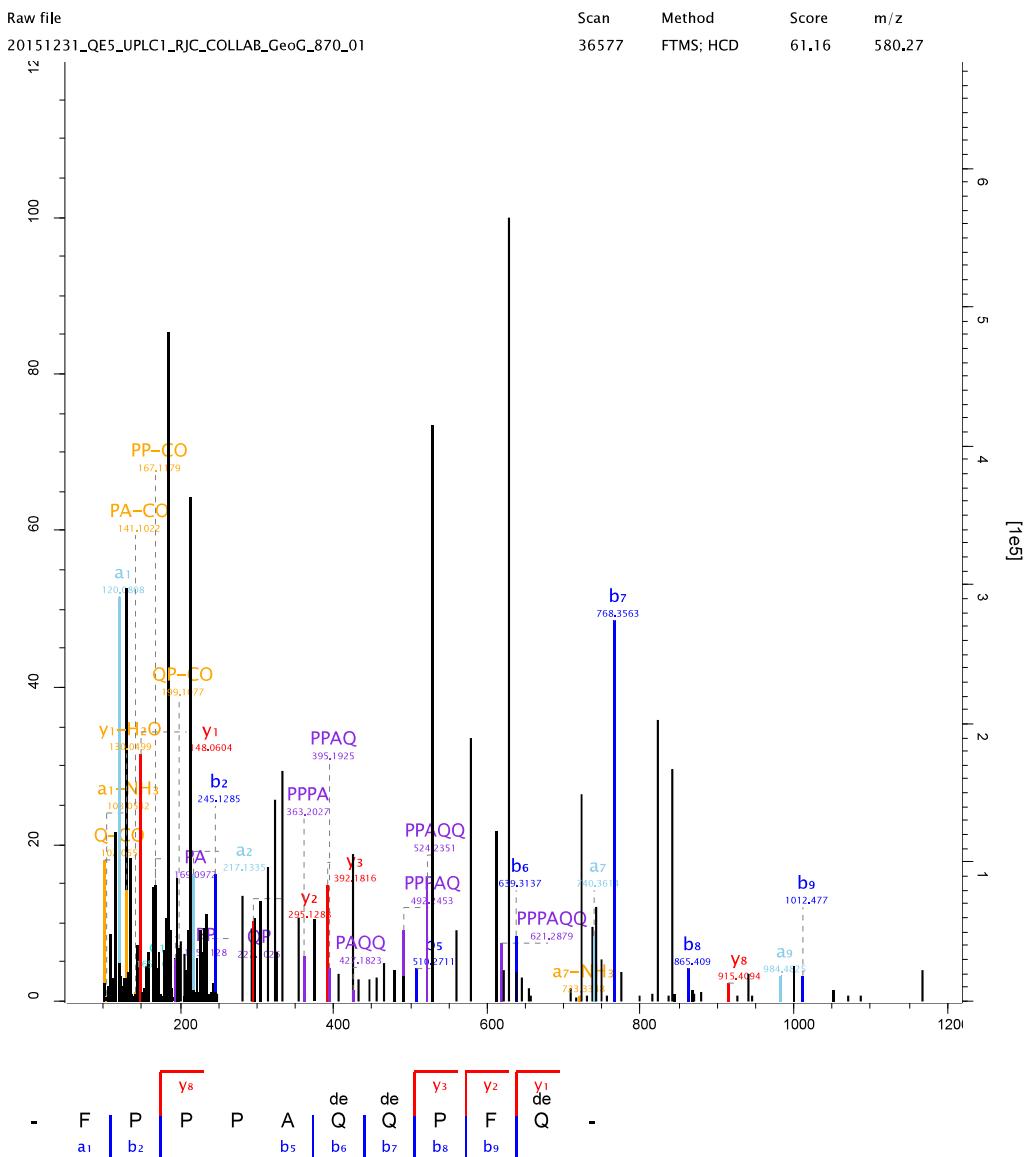
ox M P L Y₁₂ P Y₁₁ P H Y₁₀ Y₉ G Y₇ Y₆ H P G Y₁₁ Y₁₂ N Y₂ de F -

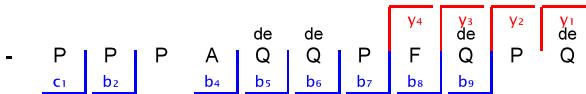
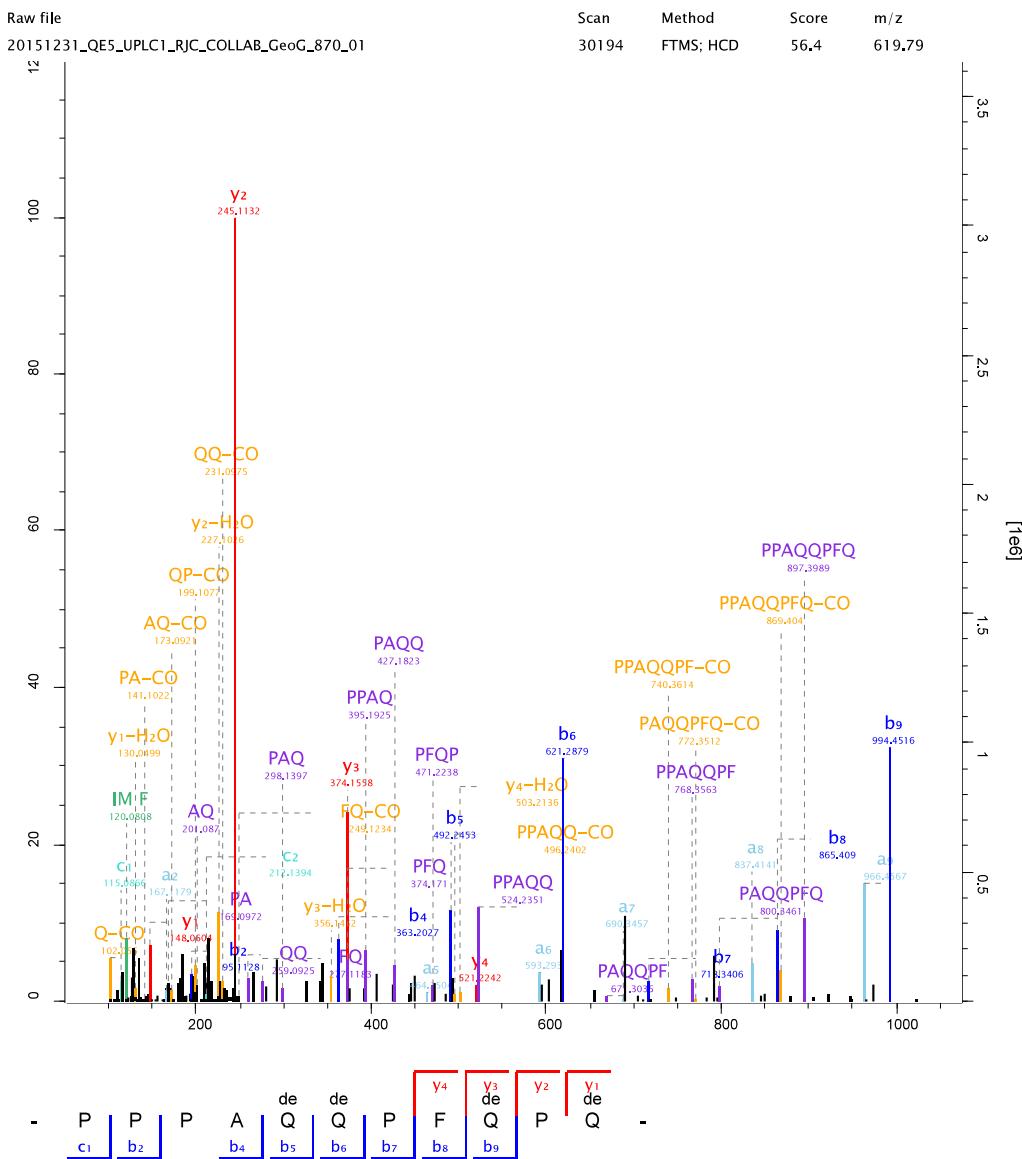


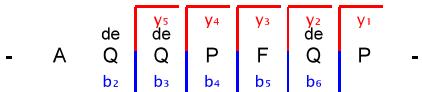
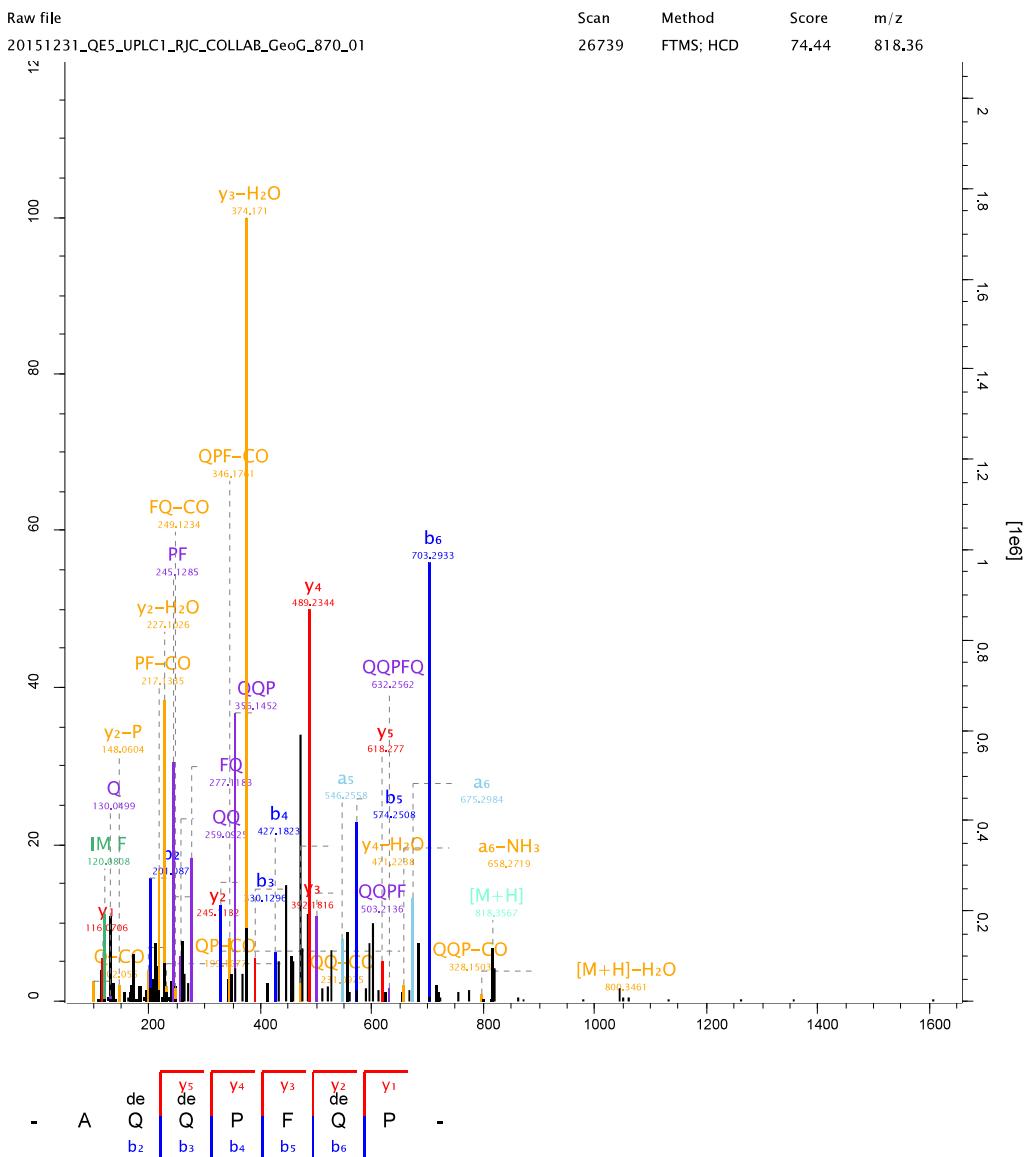


P L P P H G H

c1 b2 b3 b5 b7

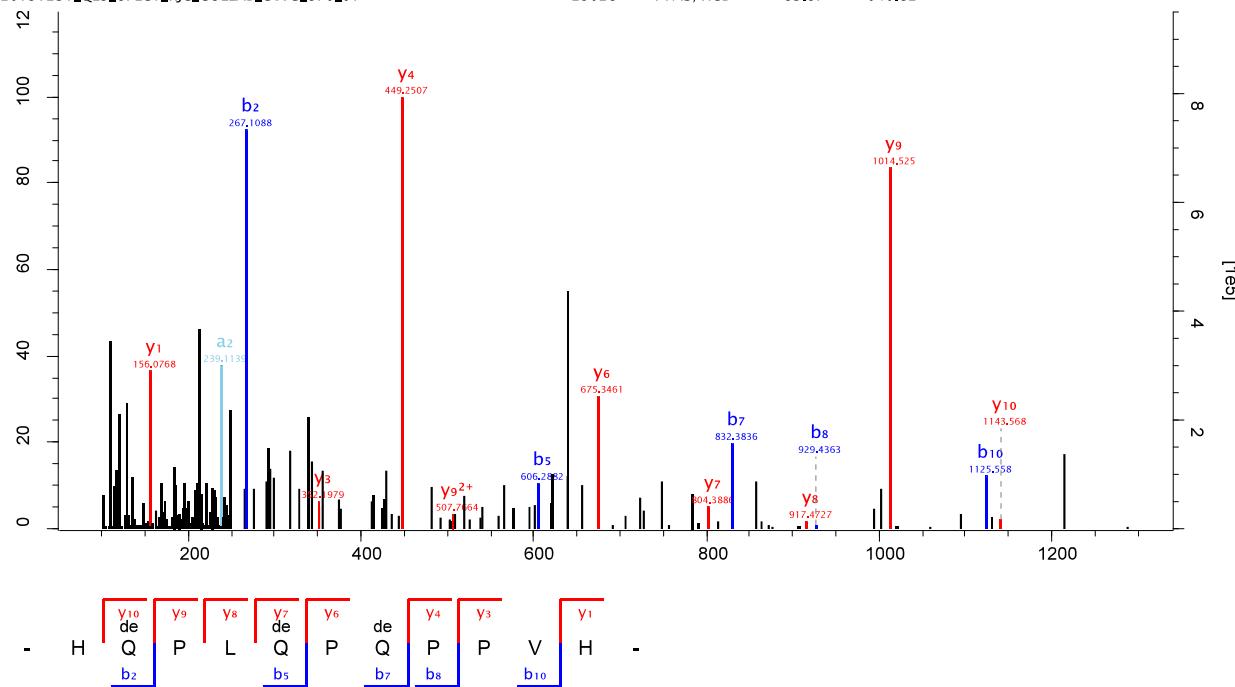


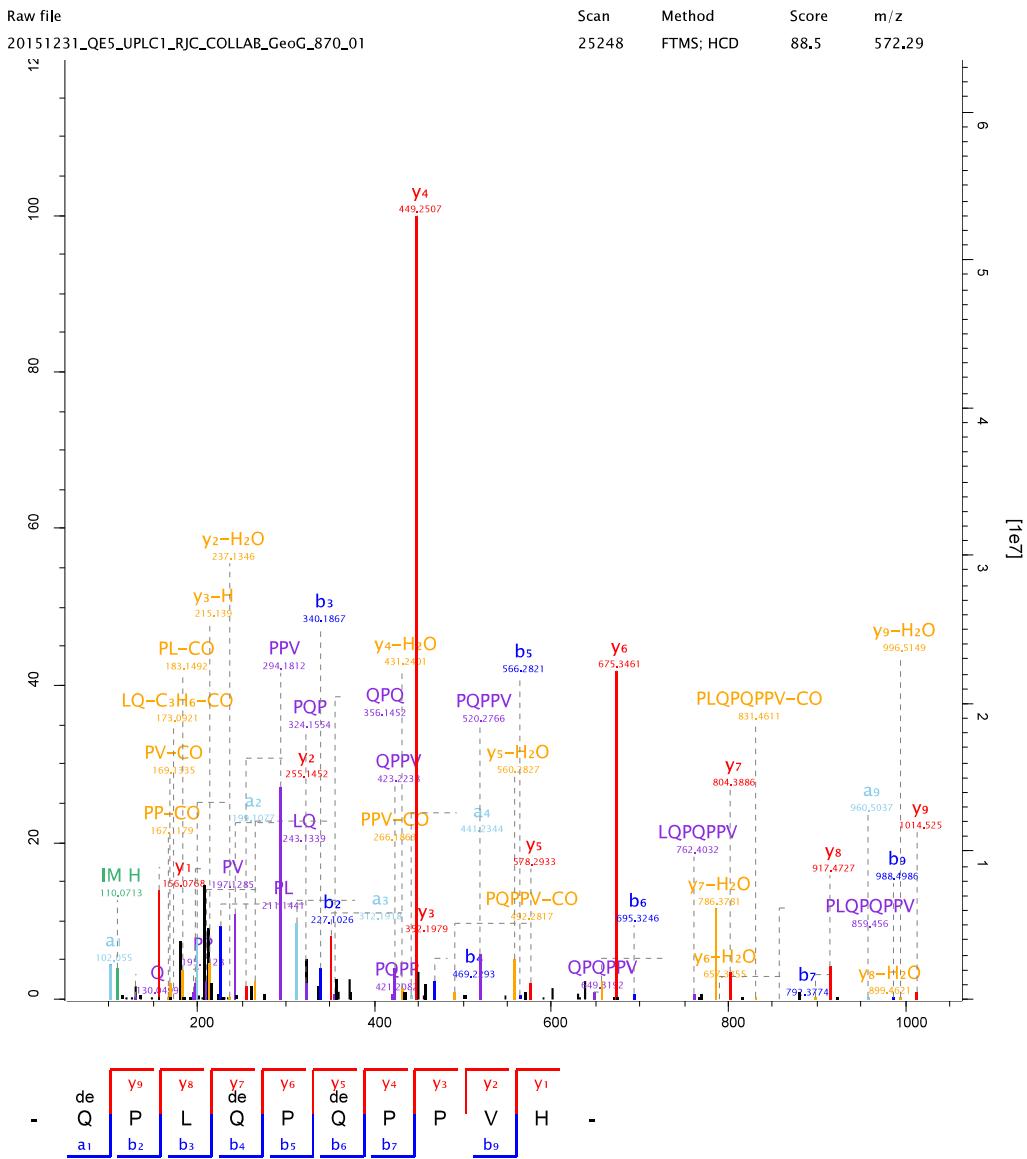


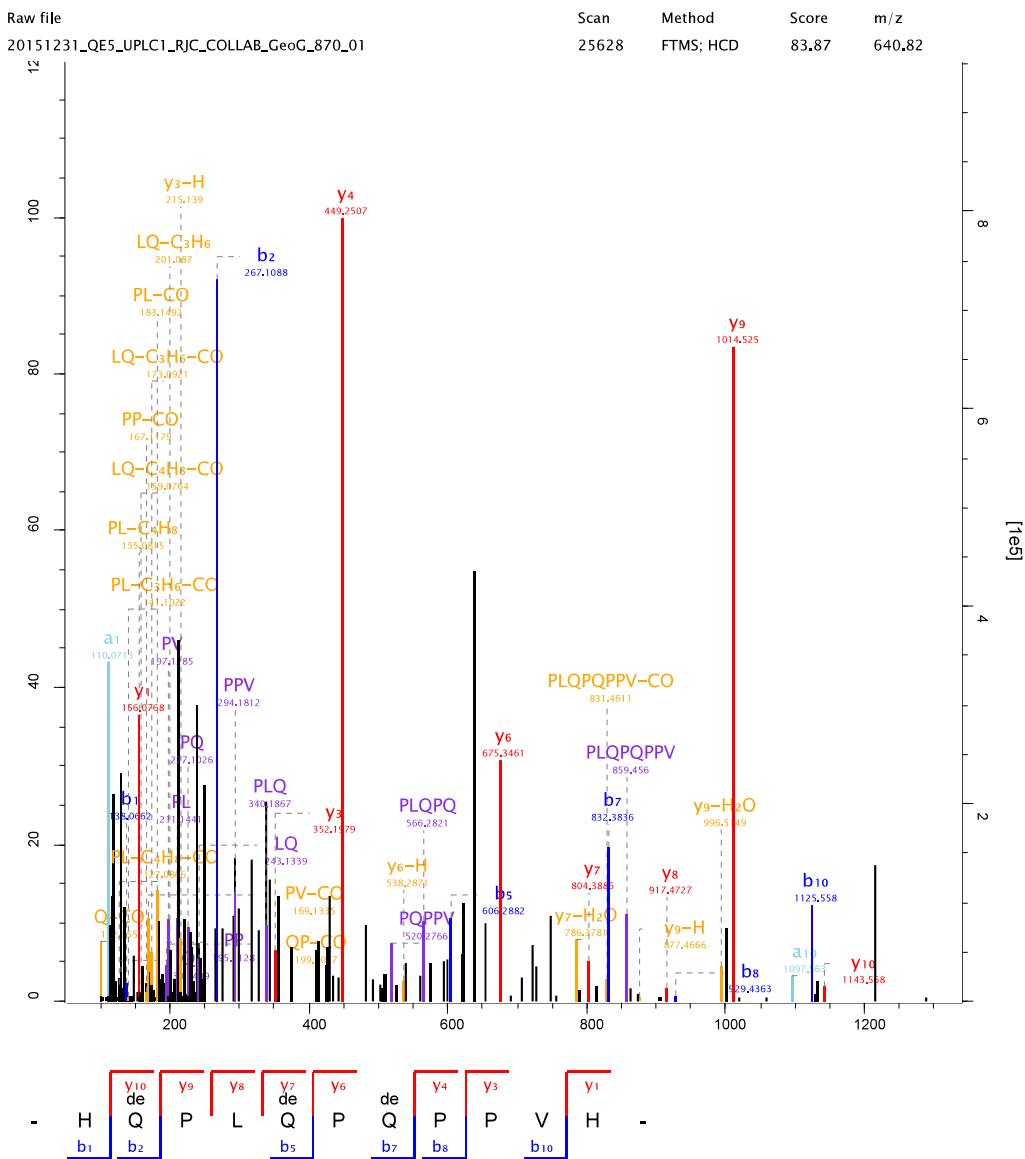


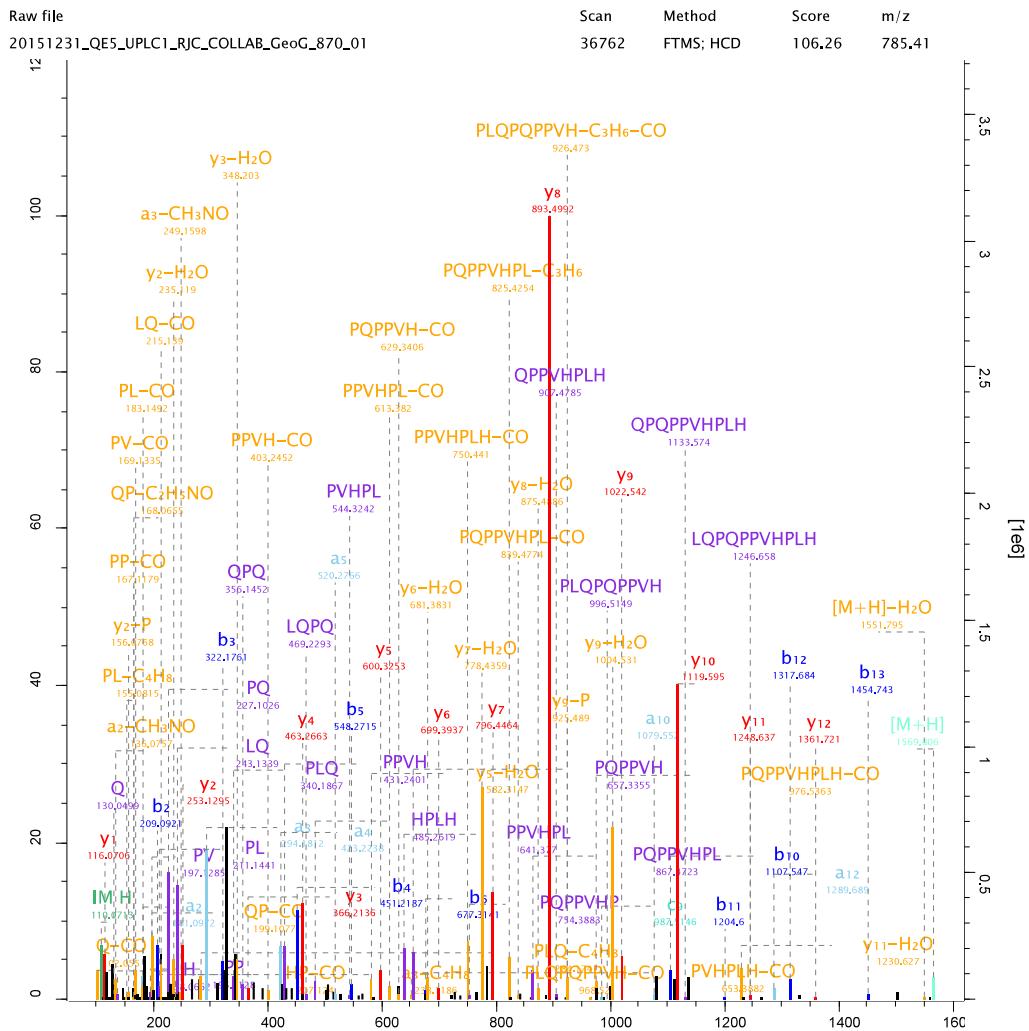
Raw file
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01

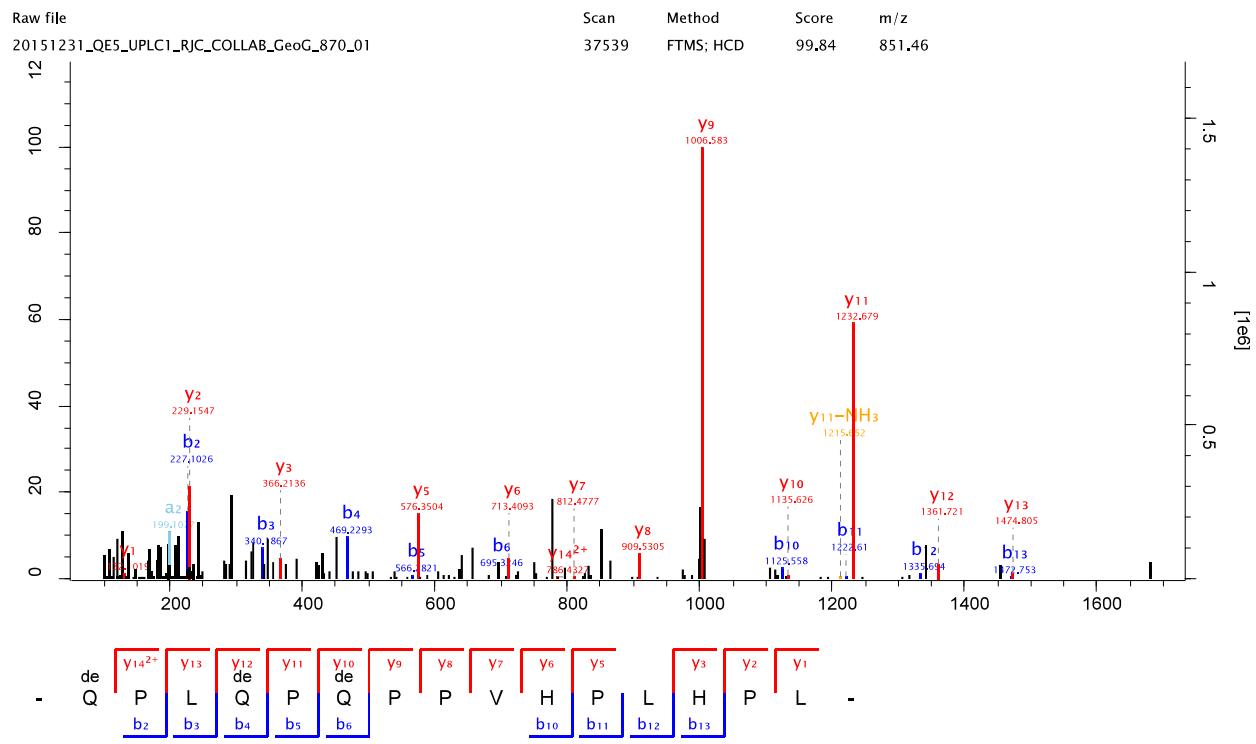
Scan 25628 Method FTMS; HCD Score 83.87 m/z 640.82

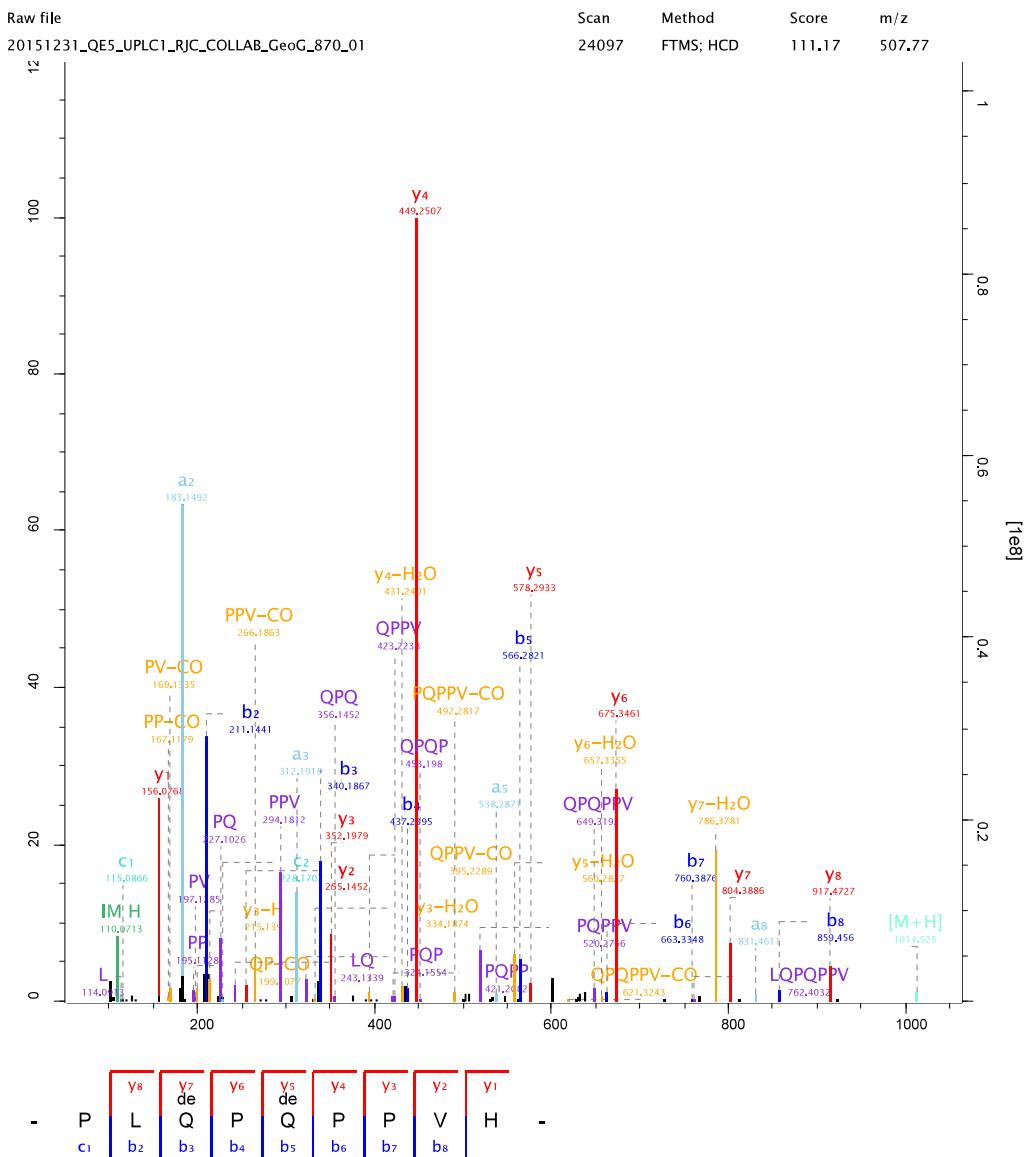




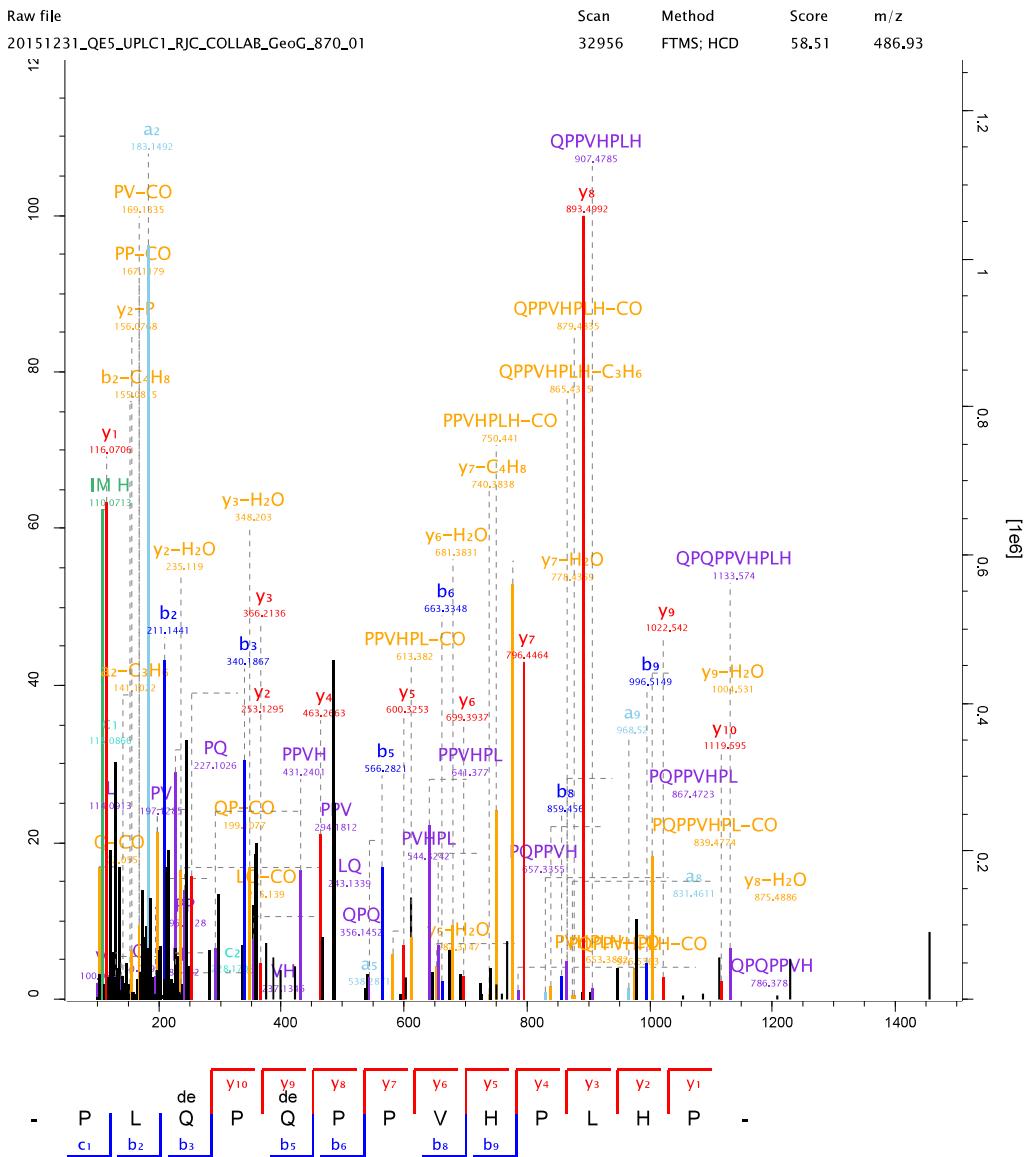




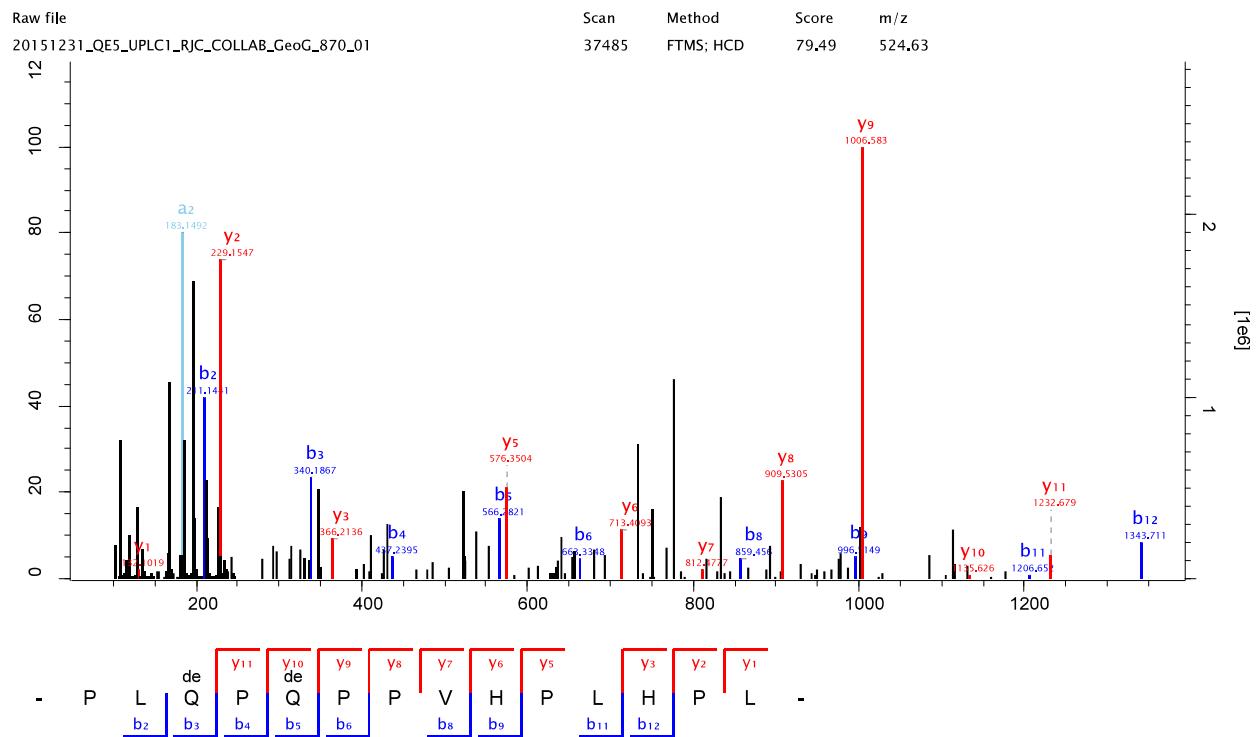


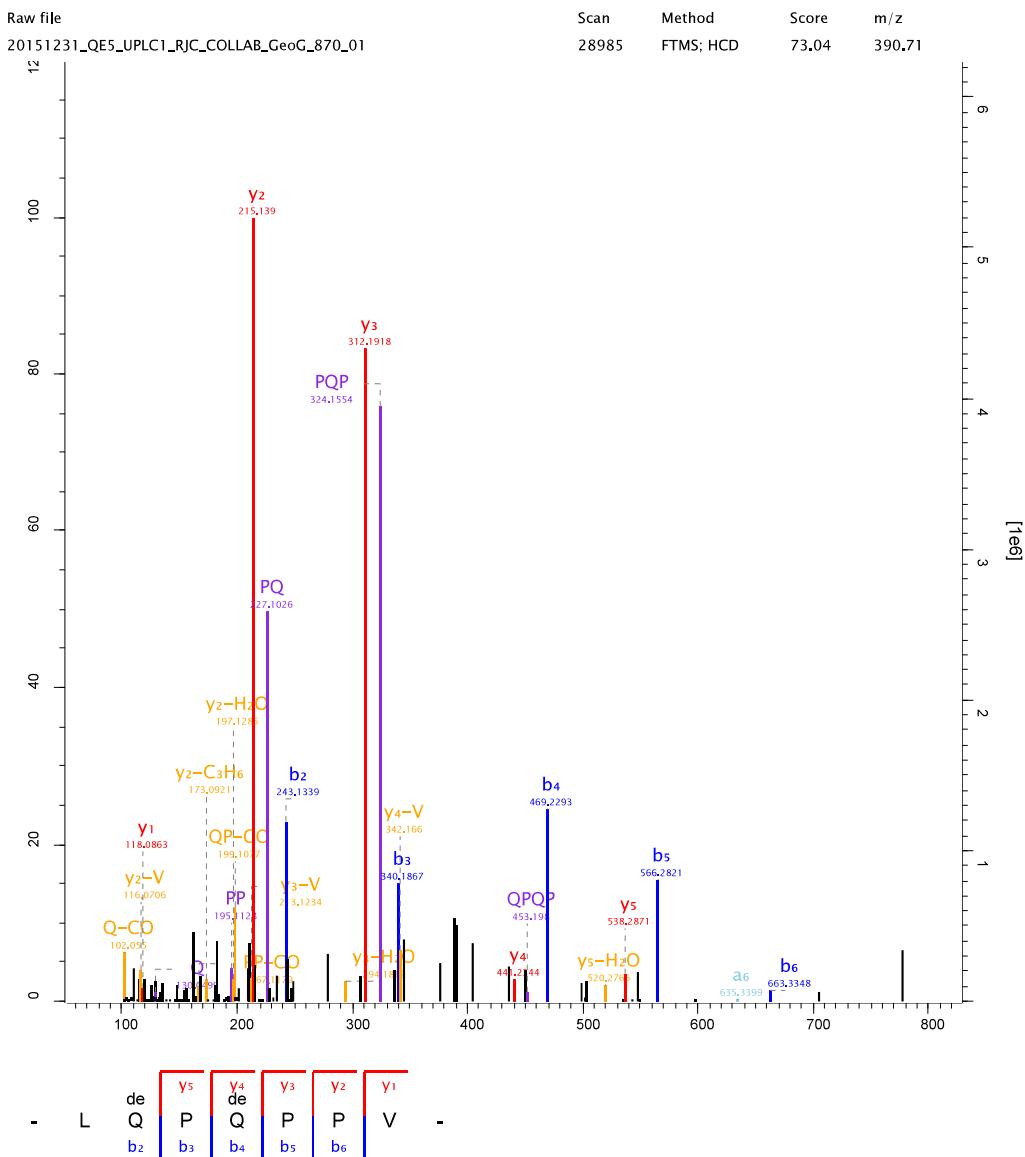


- P L Q P Q P V H -
 - c1 b2 y7 de b3 b4 b5 b6 b7 b8 y1 -



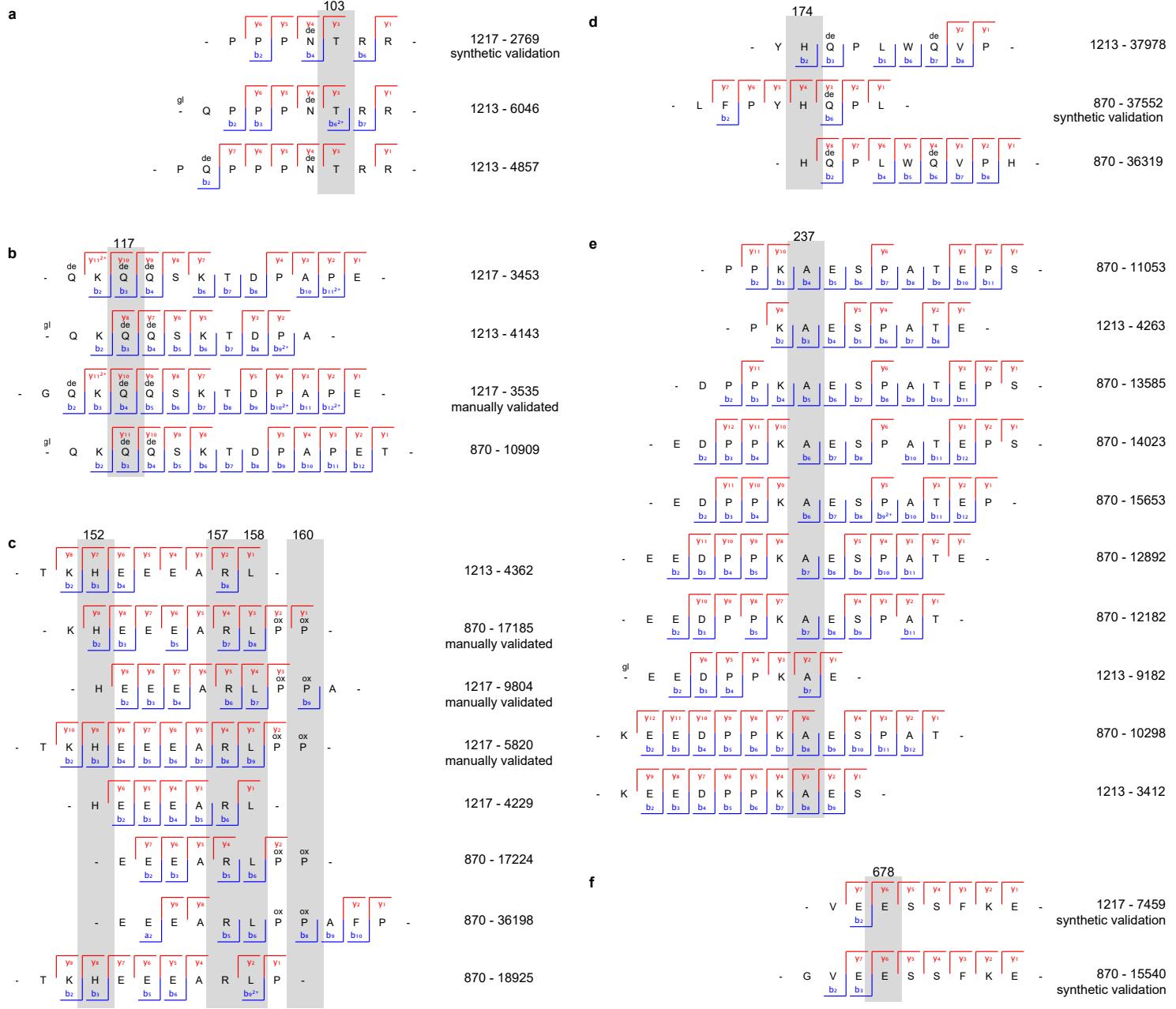
- P L de Q P y₁₀ de Q P y₈ b₅ P y₇ V b₈ H b₉ y₆ y₅ y₄ y₃ y₂ H P -



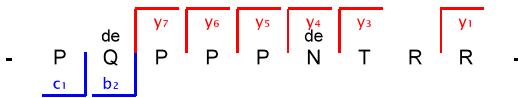
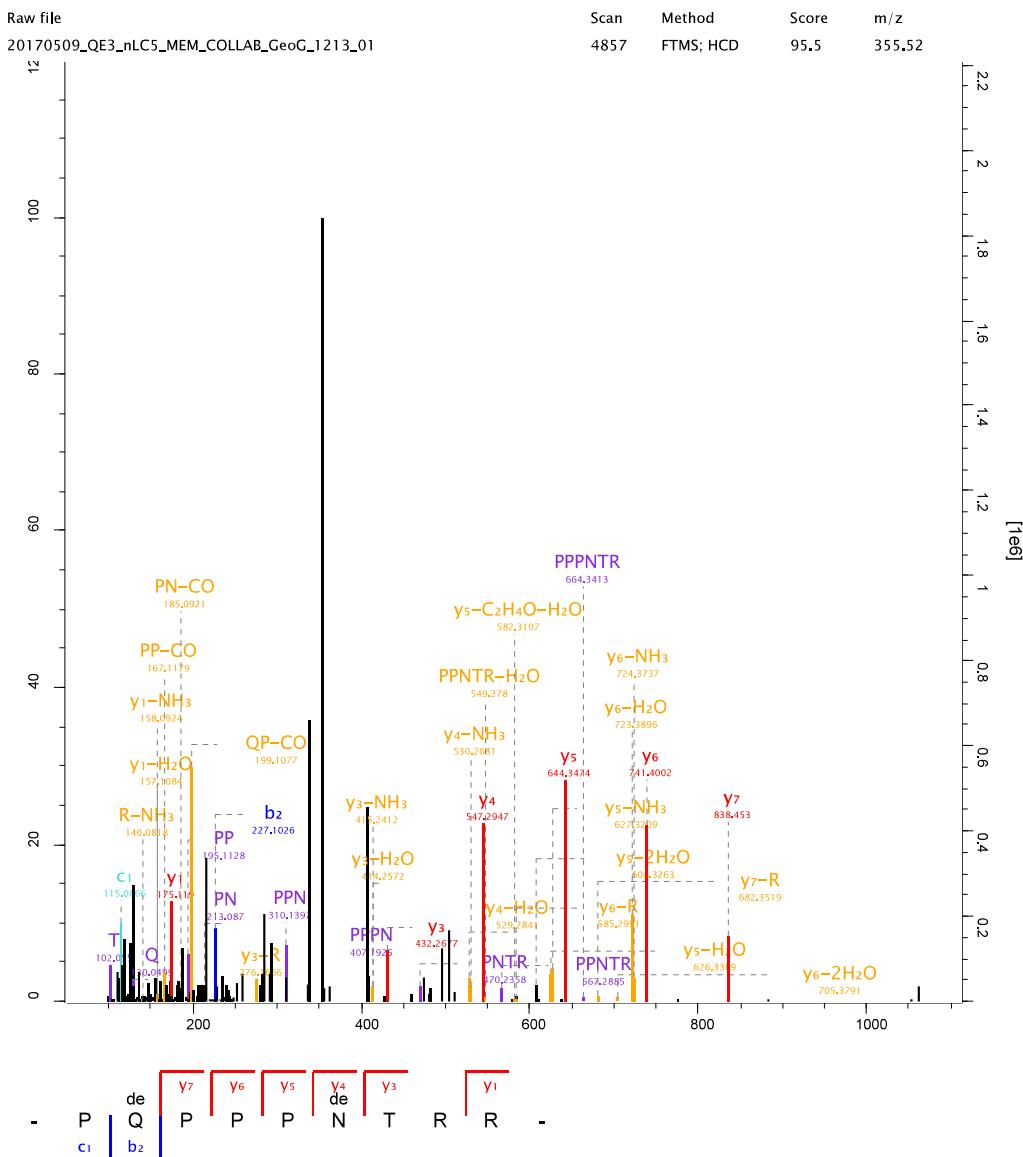


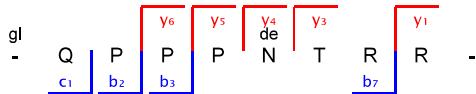
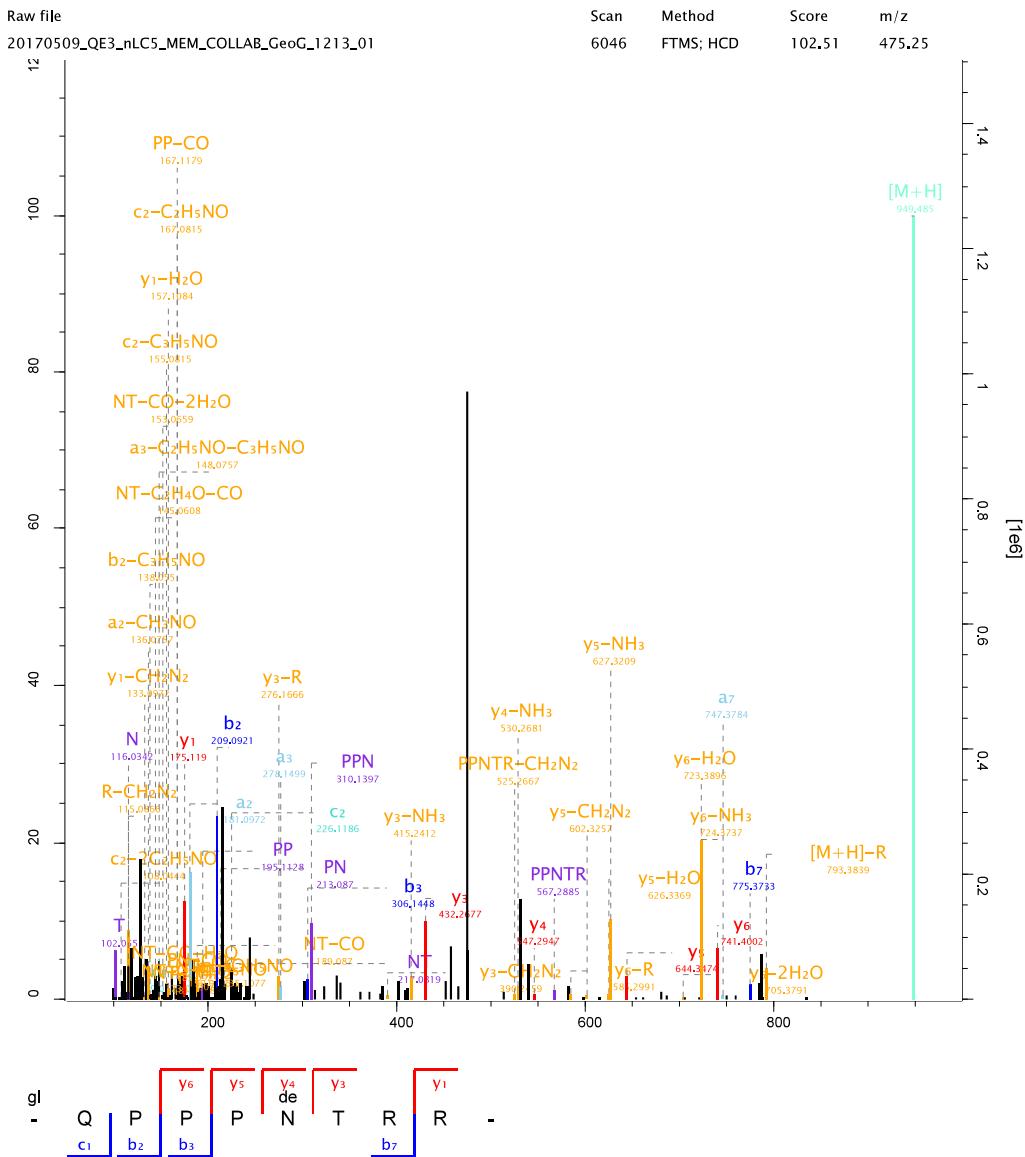
(Stagetip - scan number)

(Stagetip - scan number)



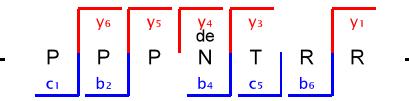
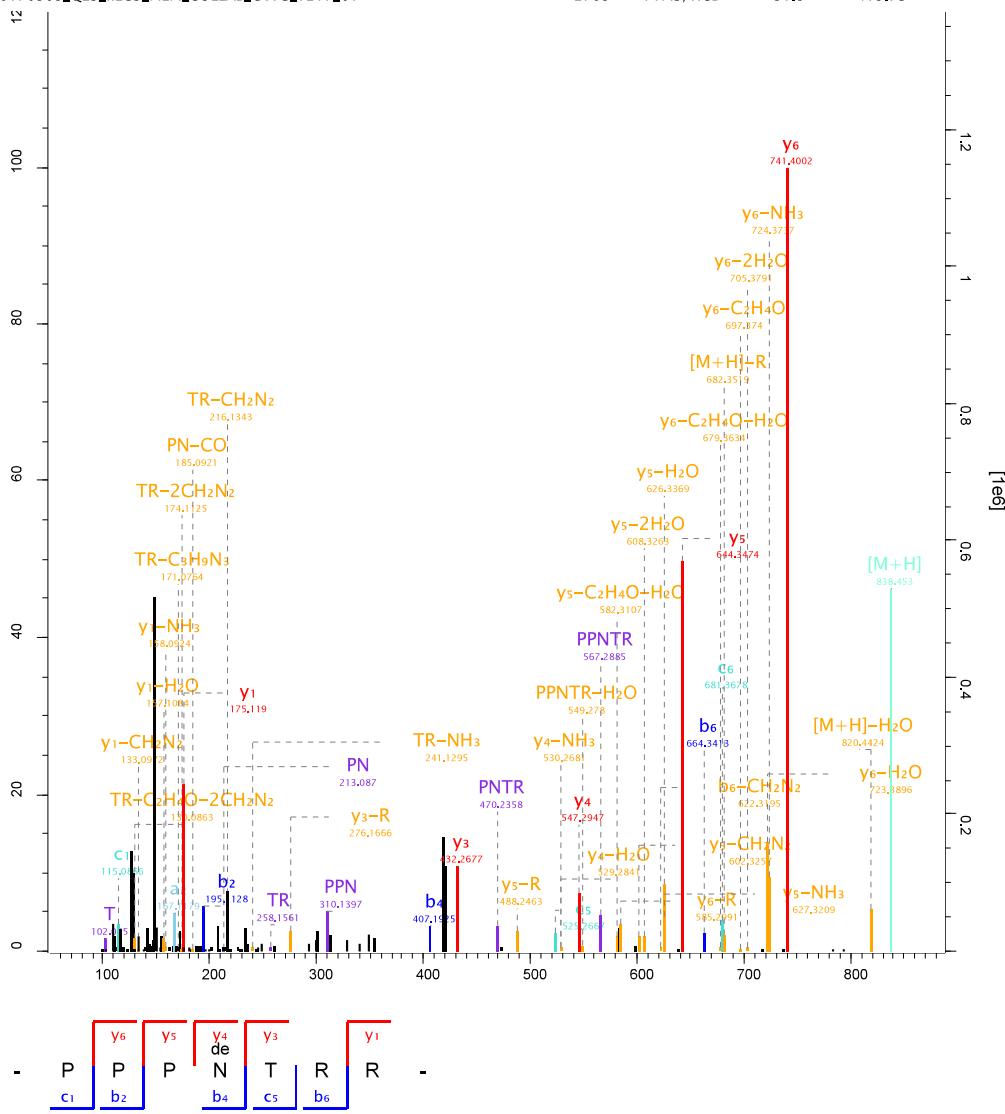
ENAM. Peptide matches to phylogenetically informative positions 103 (a), 117 (b), 152, 157, 158, and 160 (c), 174 (d), 237 (e), and 678 (f). Coordinates in reference to XP_004419201.

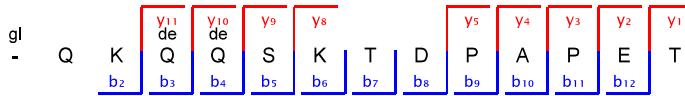
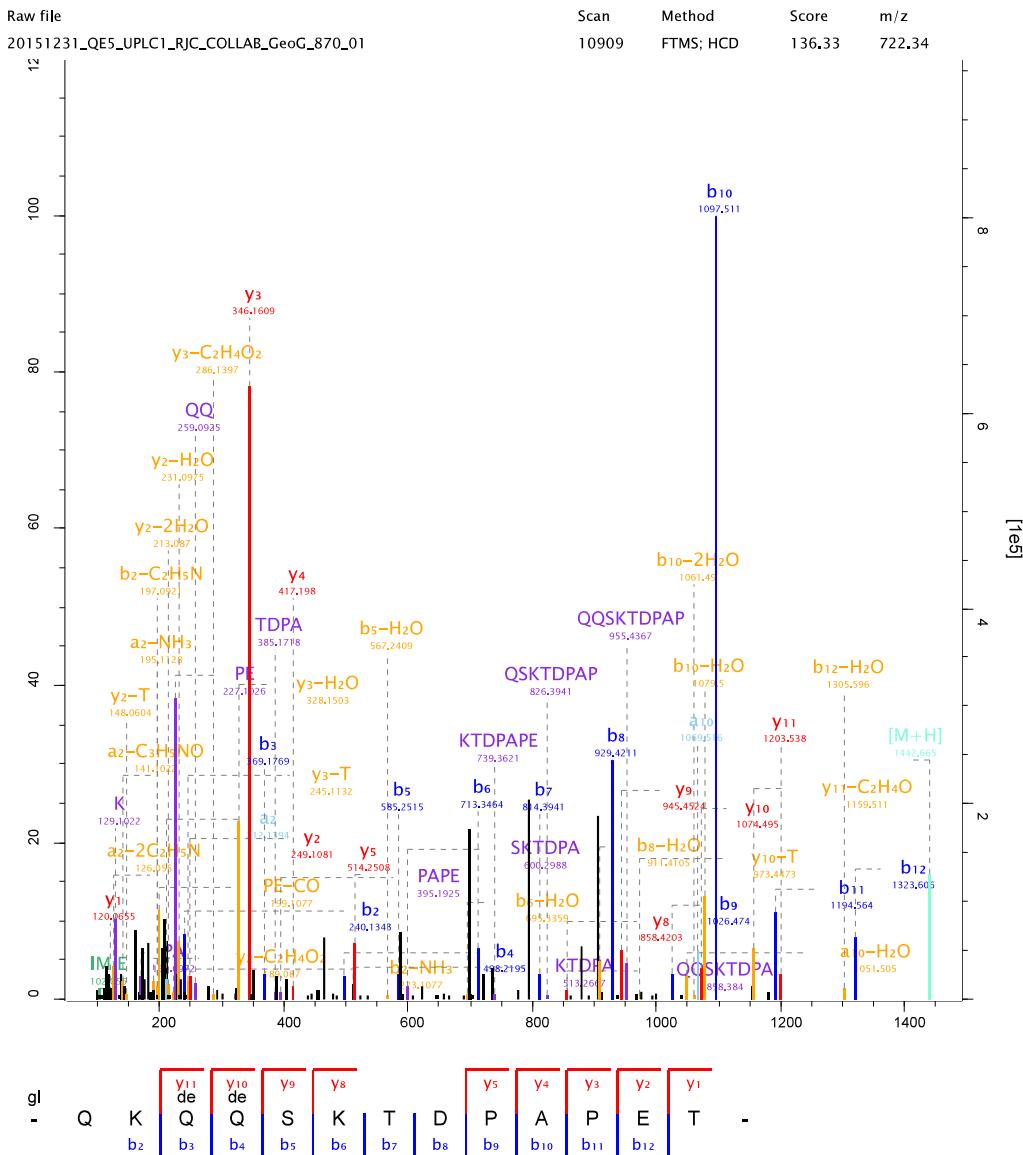




Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01

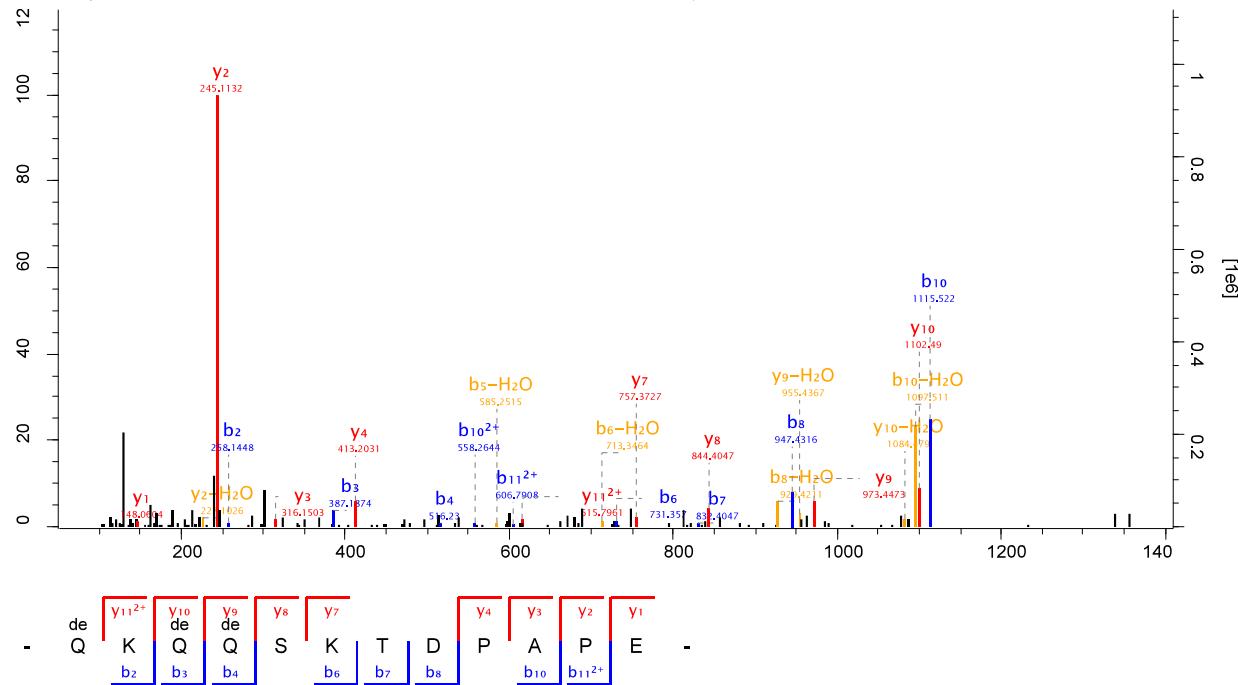
Scan 2769 Method FTMS; HCD Score 91.9 m/z 419.73





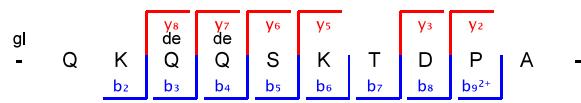
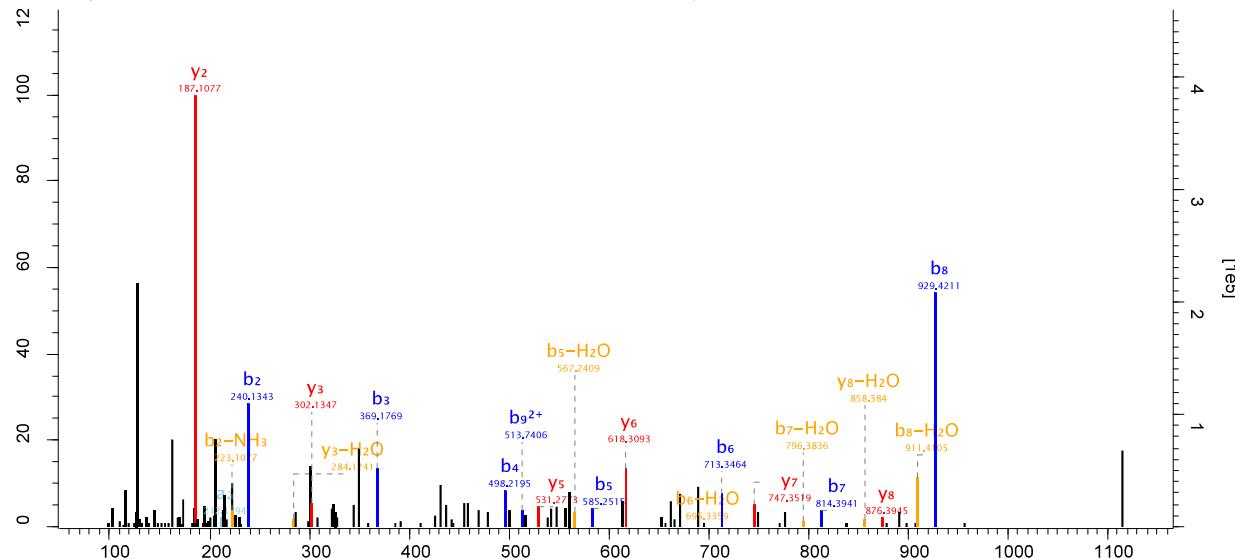
Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01

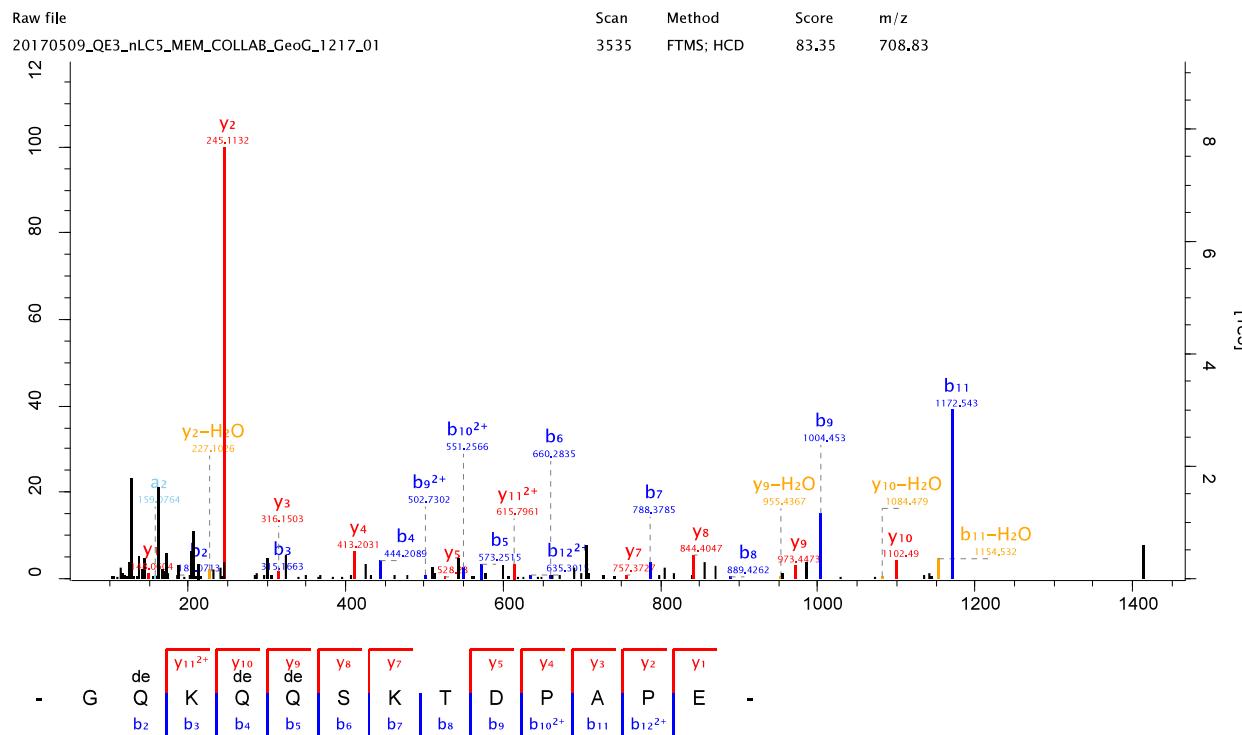
Scan 3453 Method FTMS; HCD Score 81.24 m/z 680.32



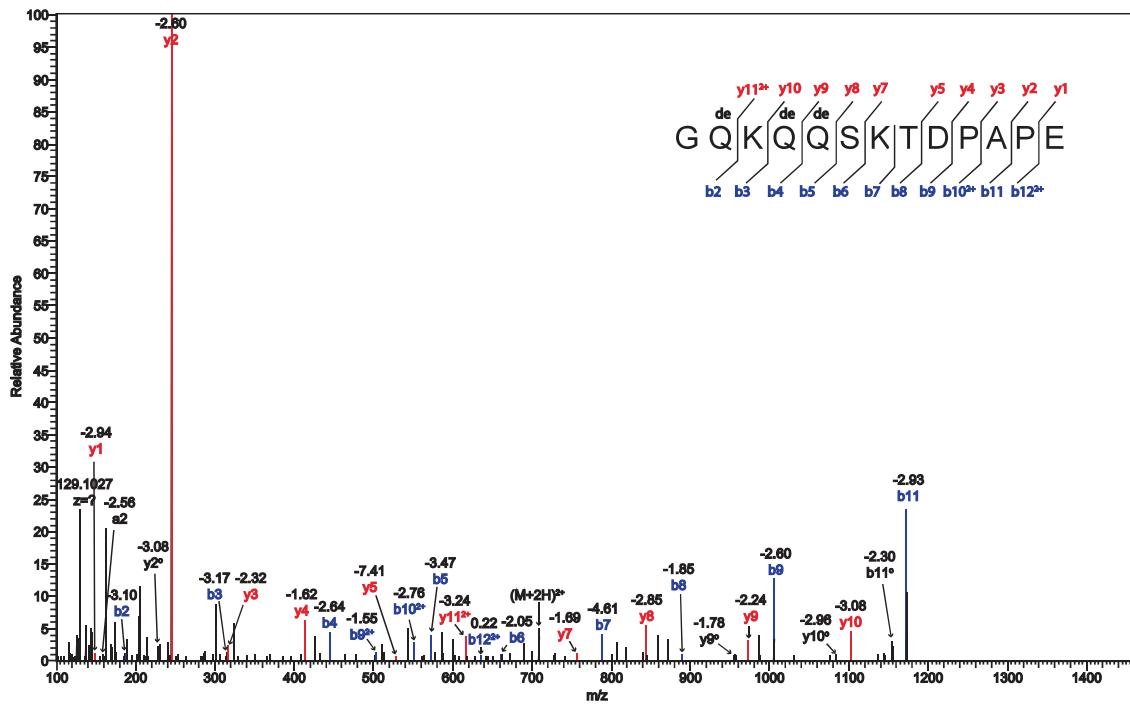
Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1213_01

Scan 4143 Method FTMS; HCD Score 83.77 m/z 558.26



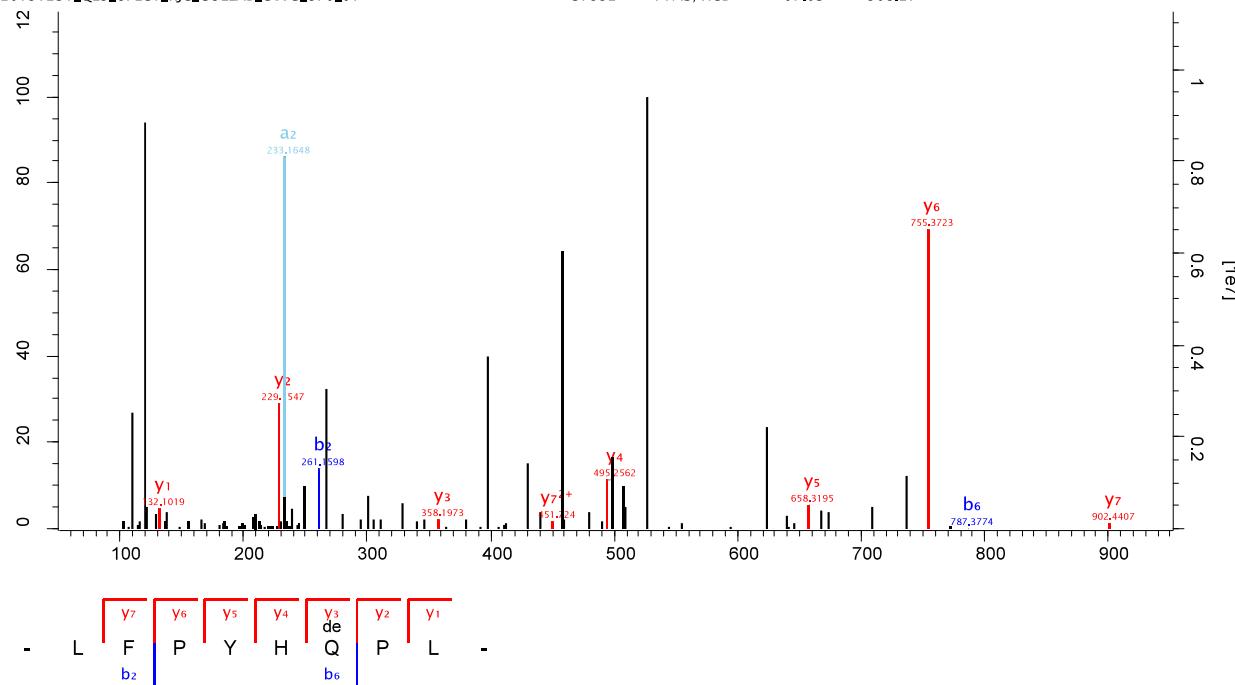


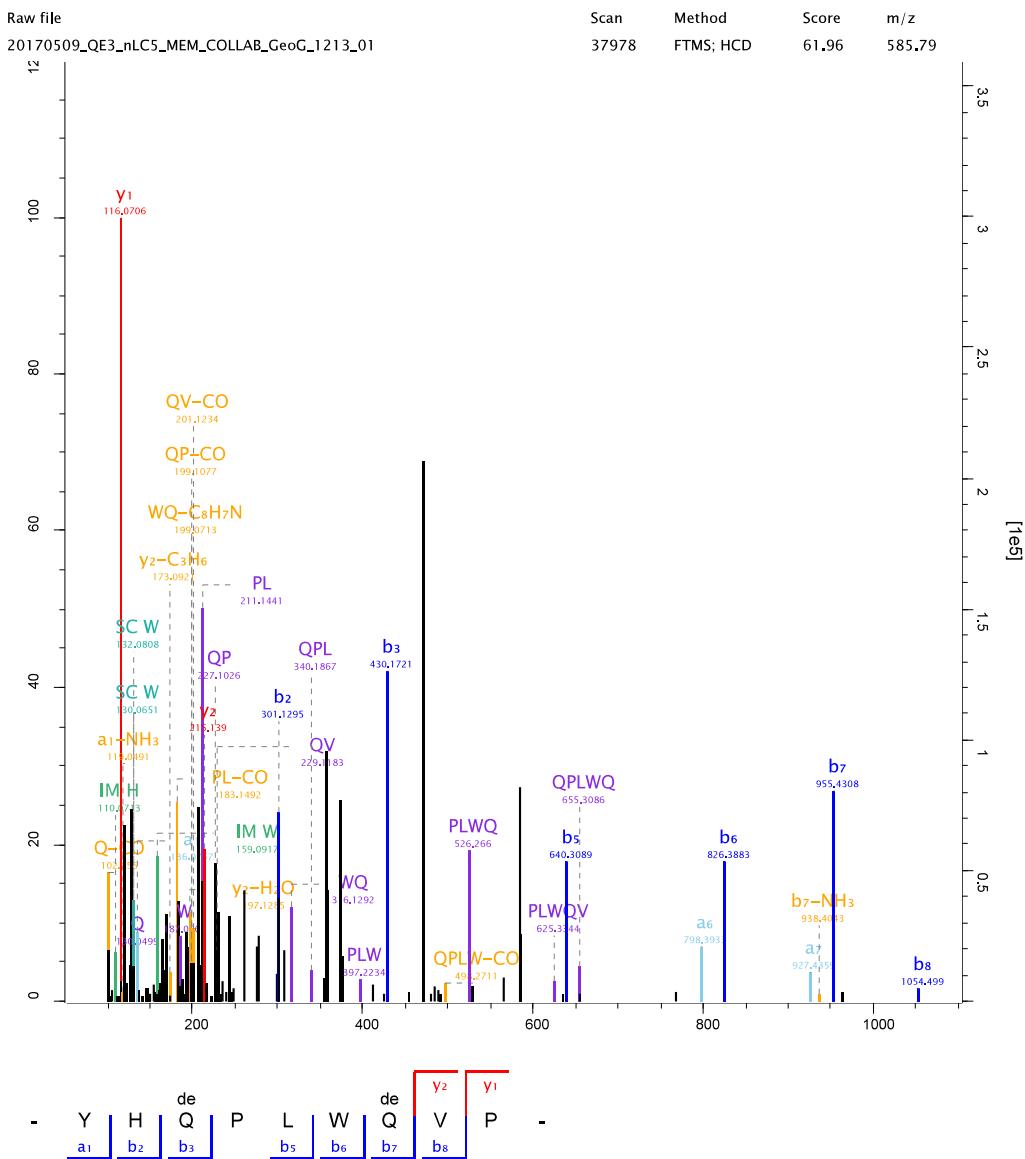
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01 #3535 RT: 15.81 AV: 1 NL: 1.91E5
T: FTMS + p NSI d Full ms2 708.83@hcd28.00 [100.00-1465.00]

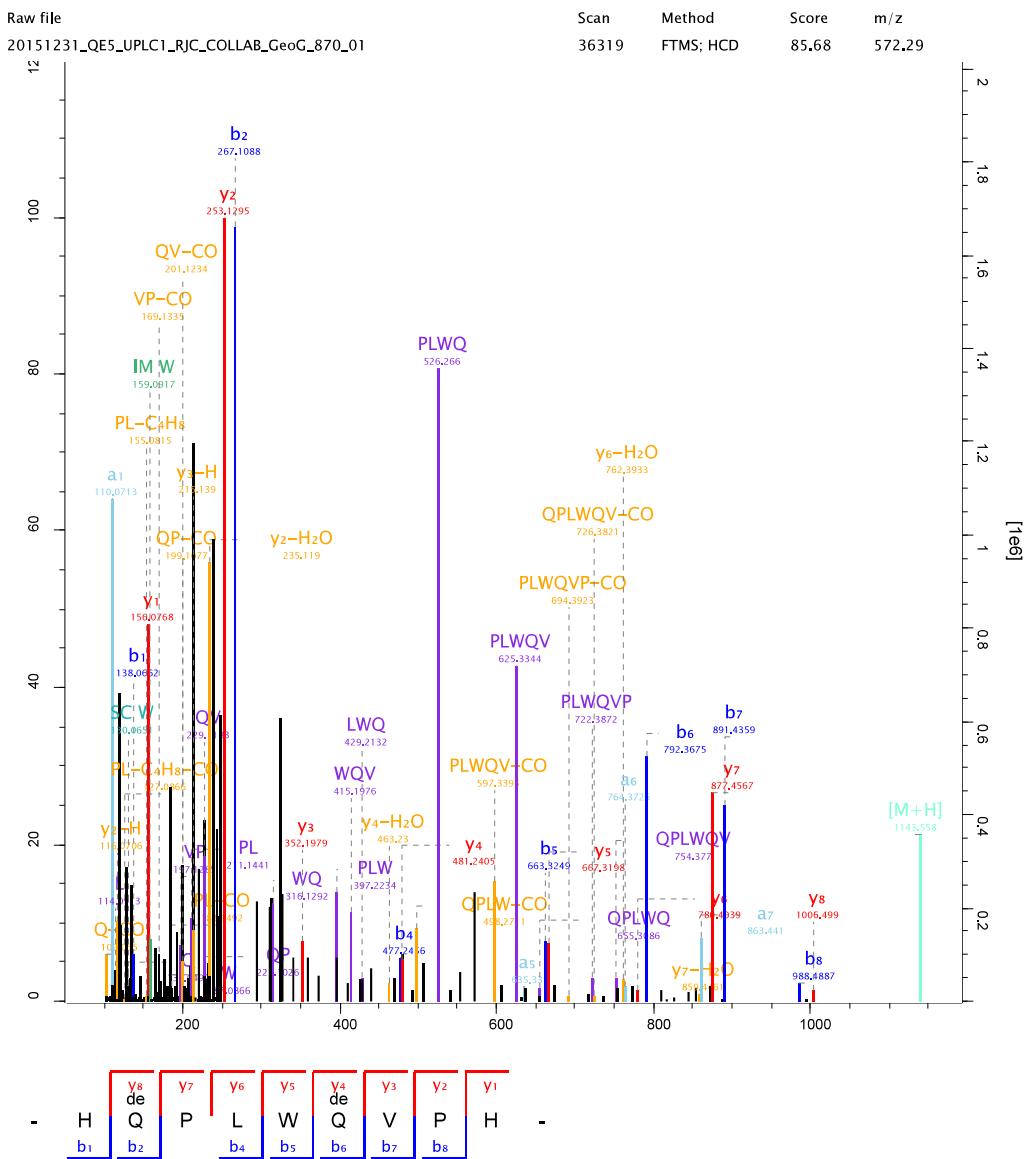


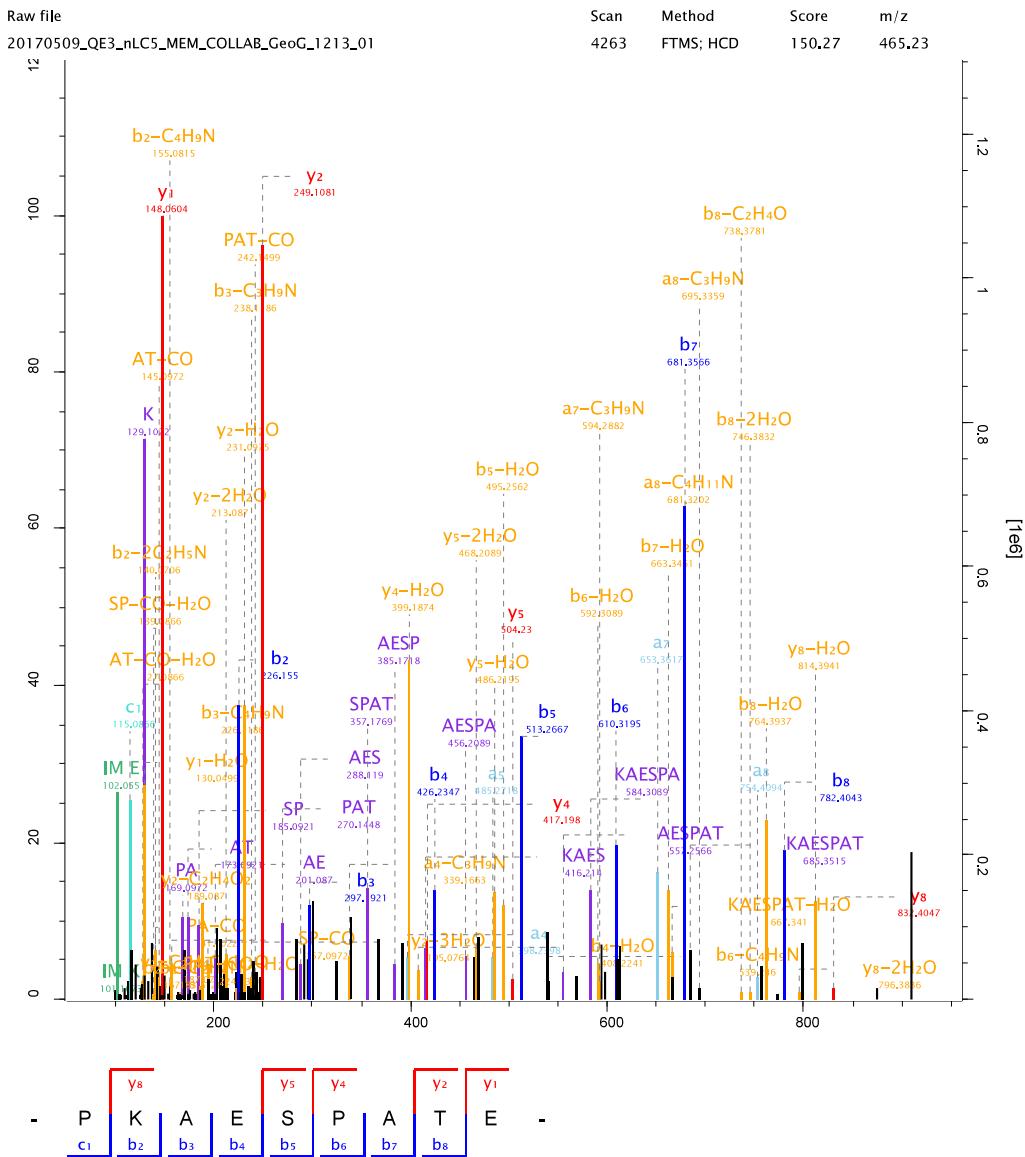
Raw file
20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01

Scan 37552 Method FTMS; HCD Score 67.03 m/z 508.27

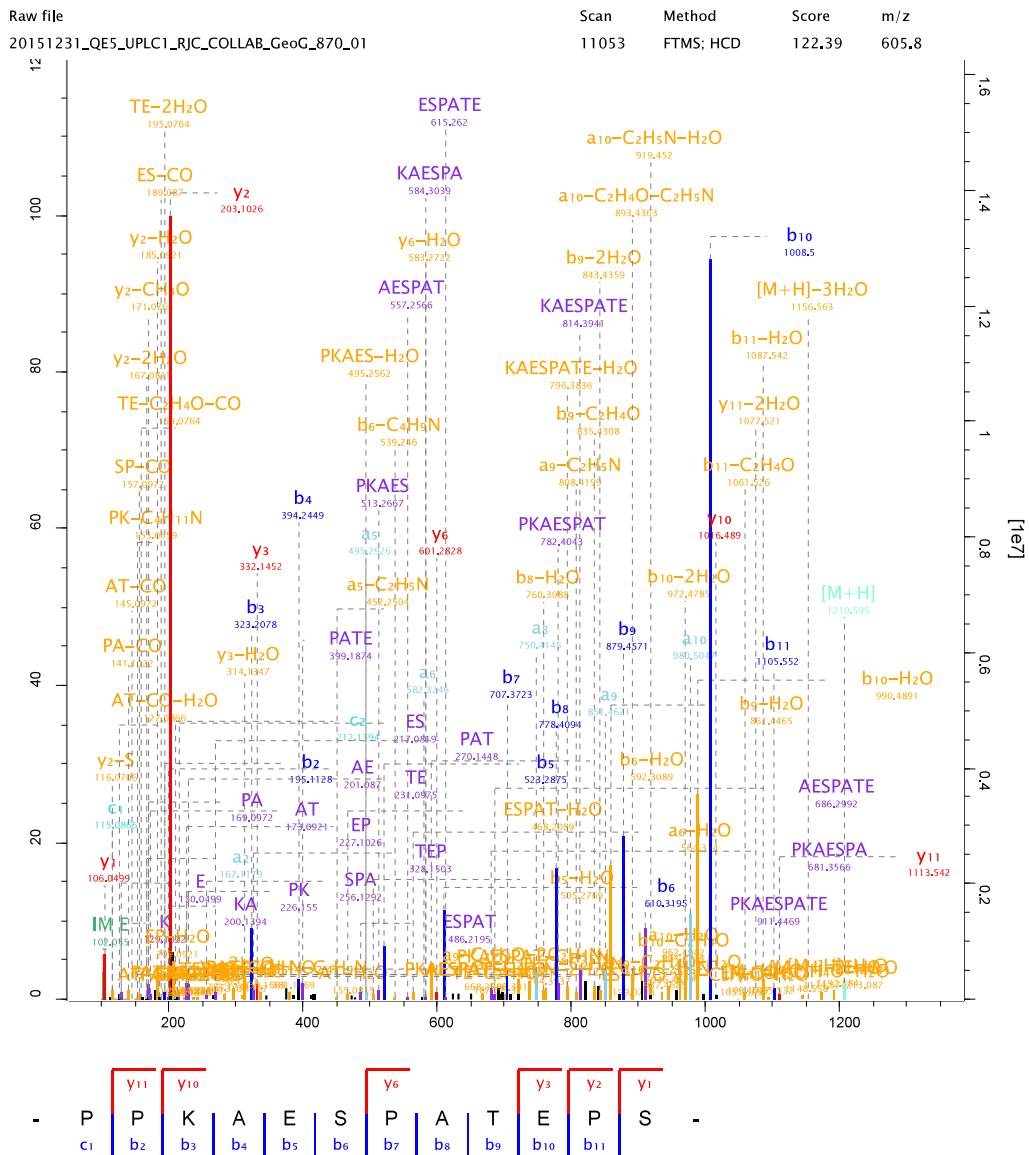


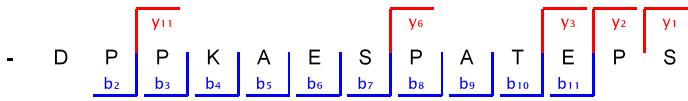
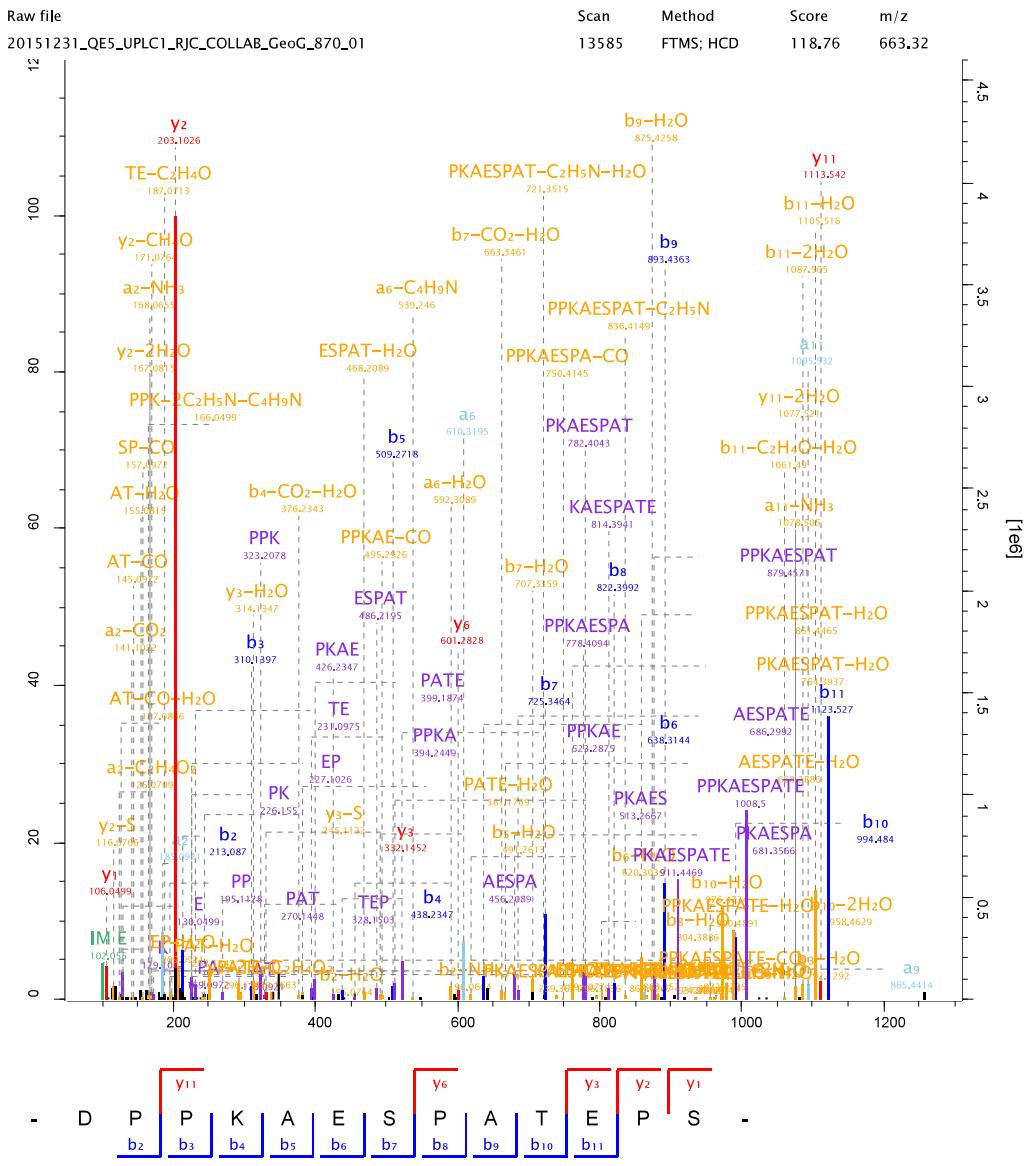


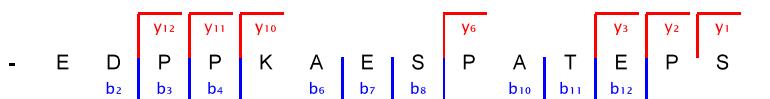
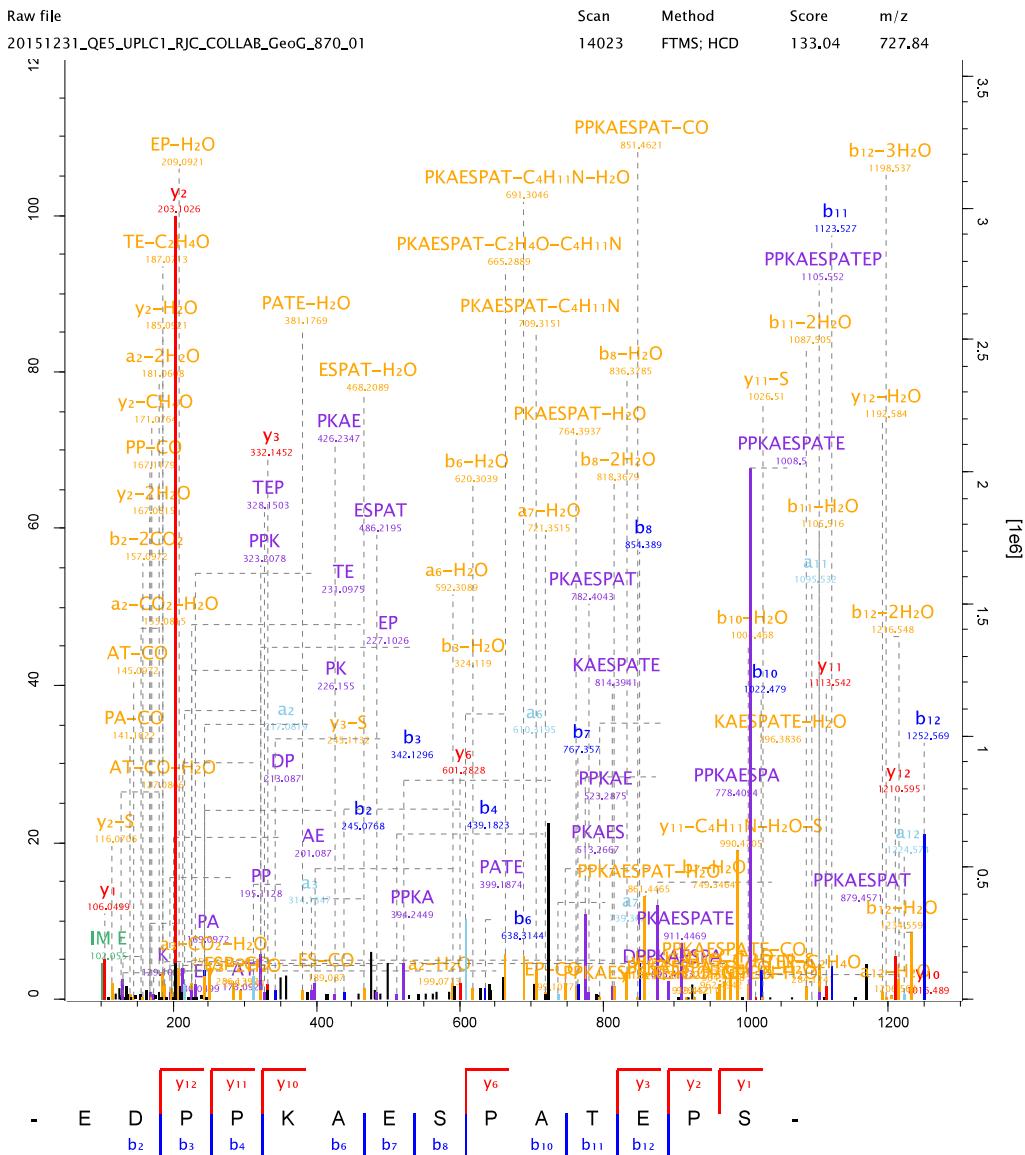


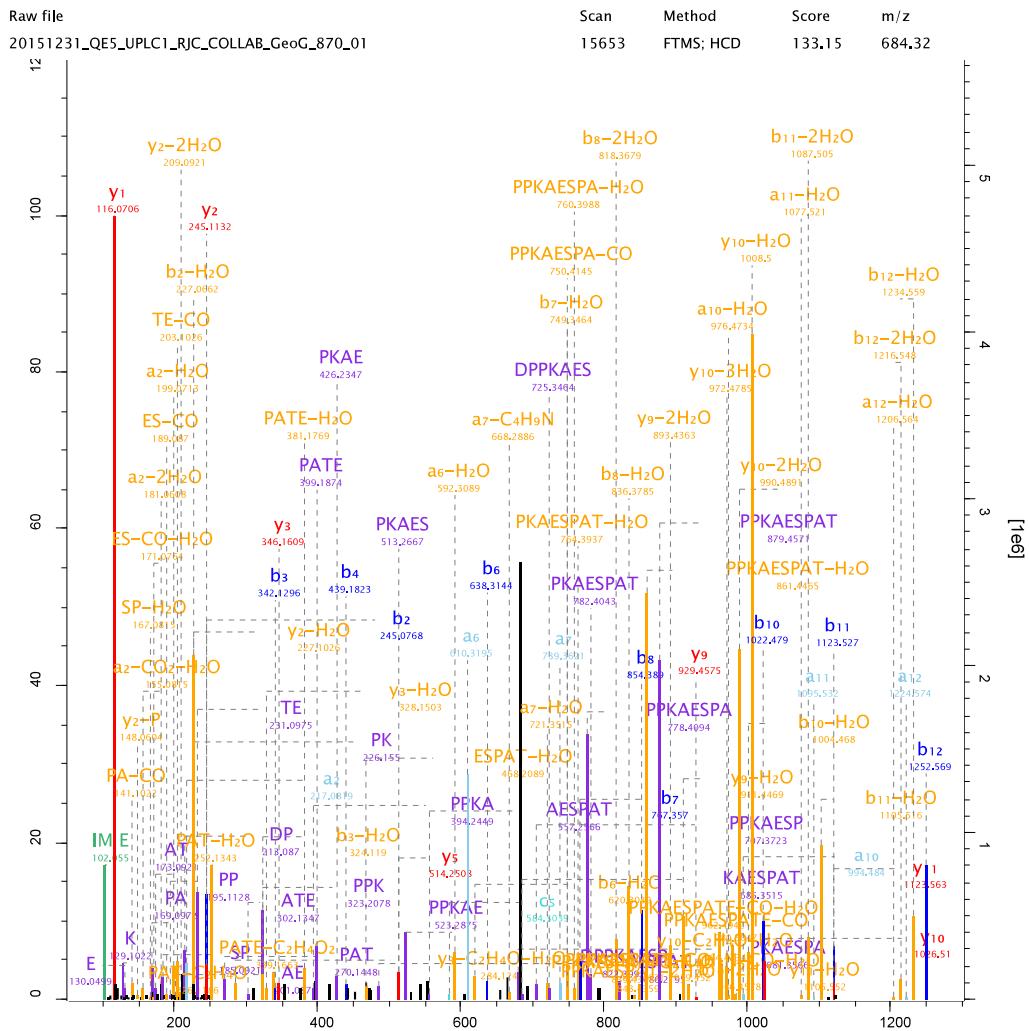


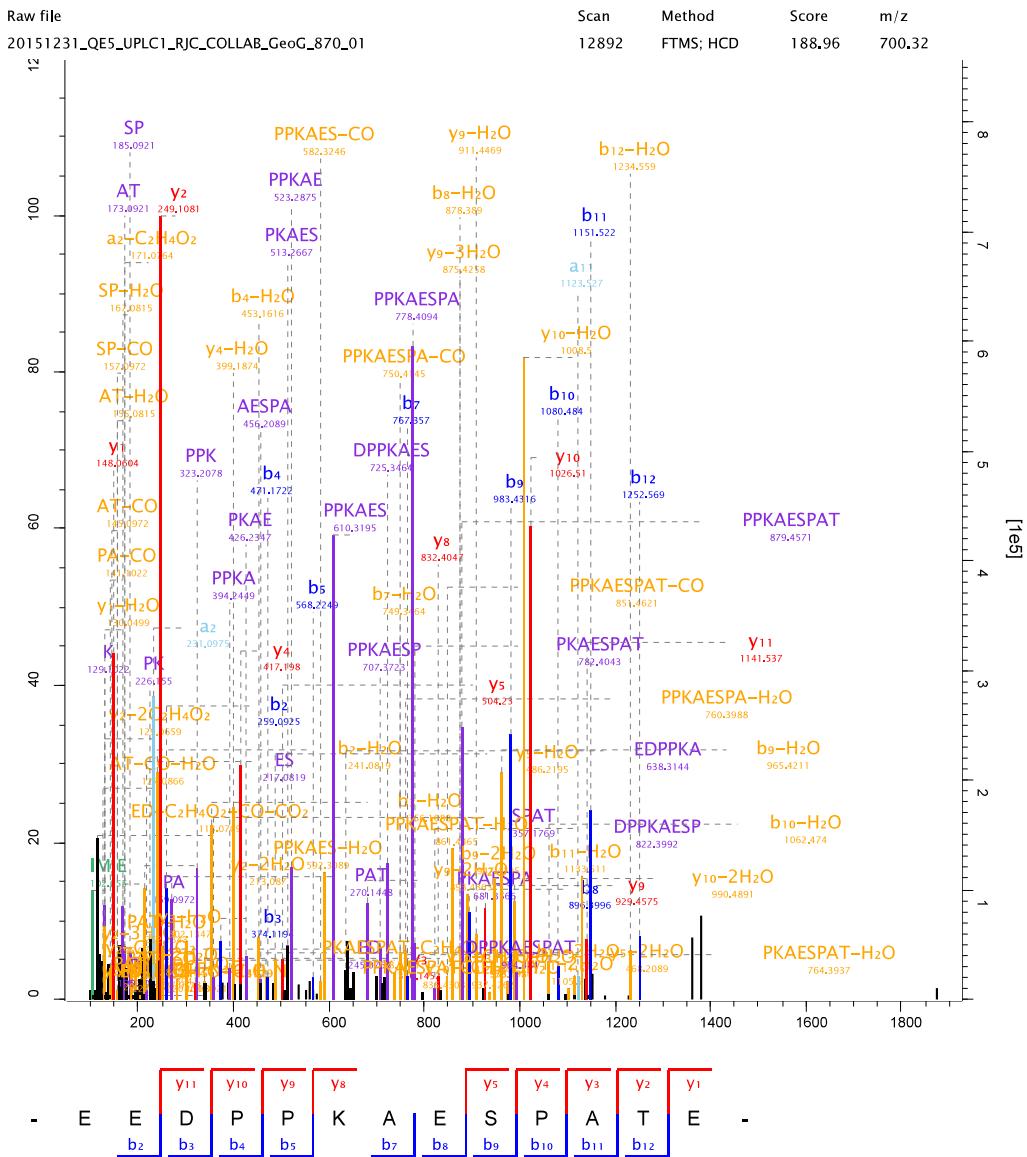
- P y₈ K b₂ A b₃ E b₄ S y₅ P b₅ A b₆ T b₇ E b₈ -







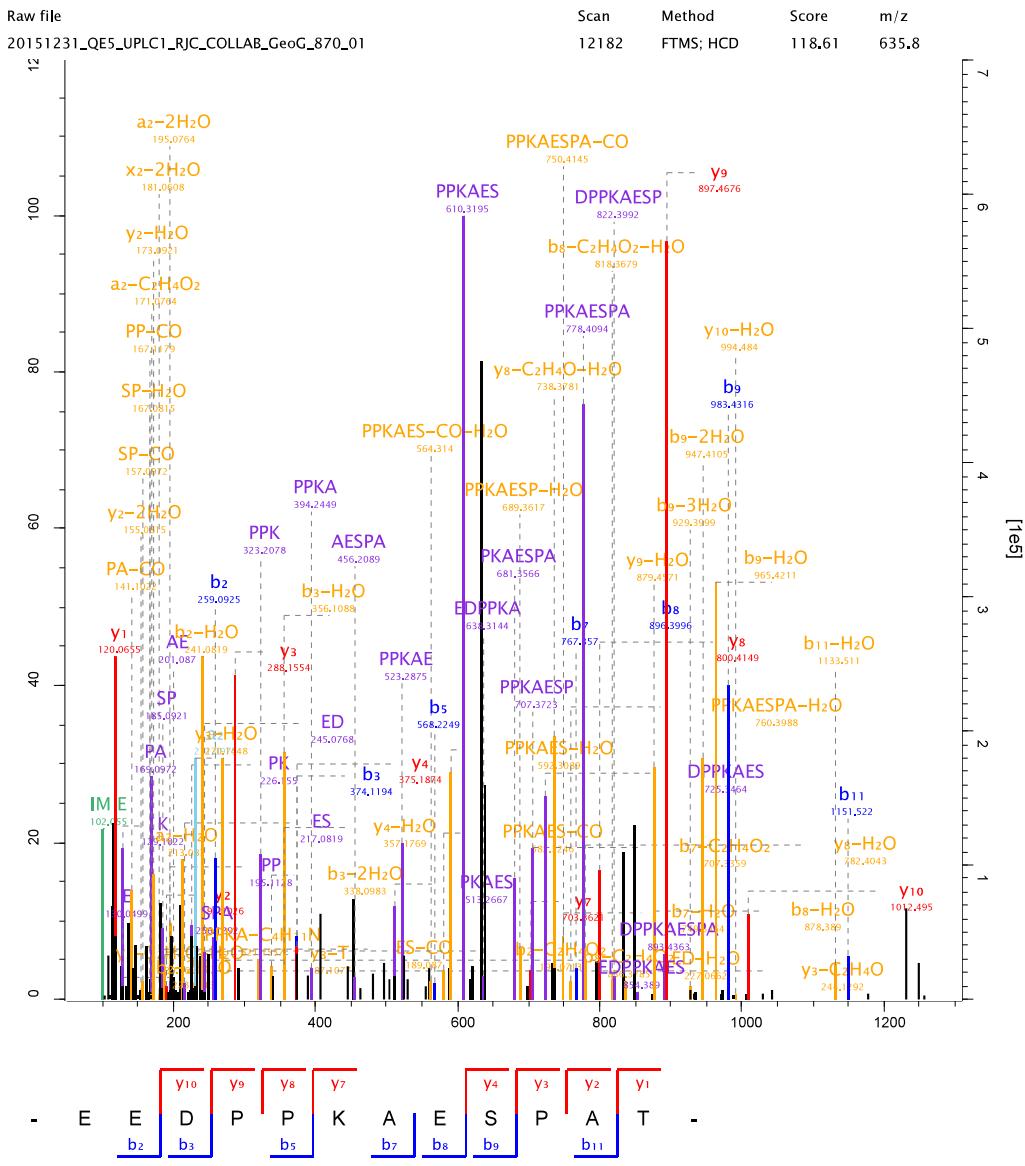




E E D P P K A E S P A T E

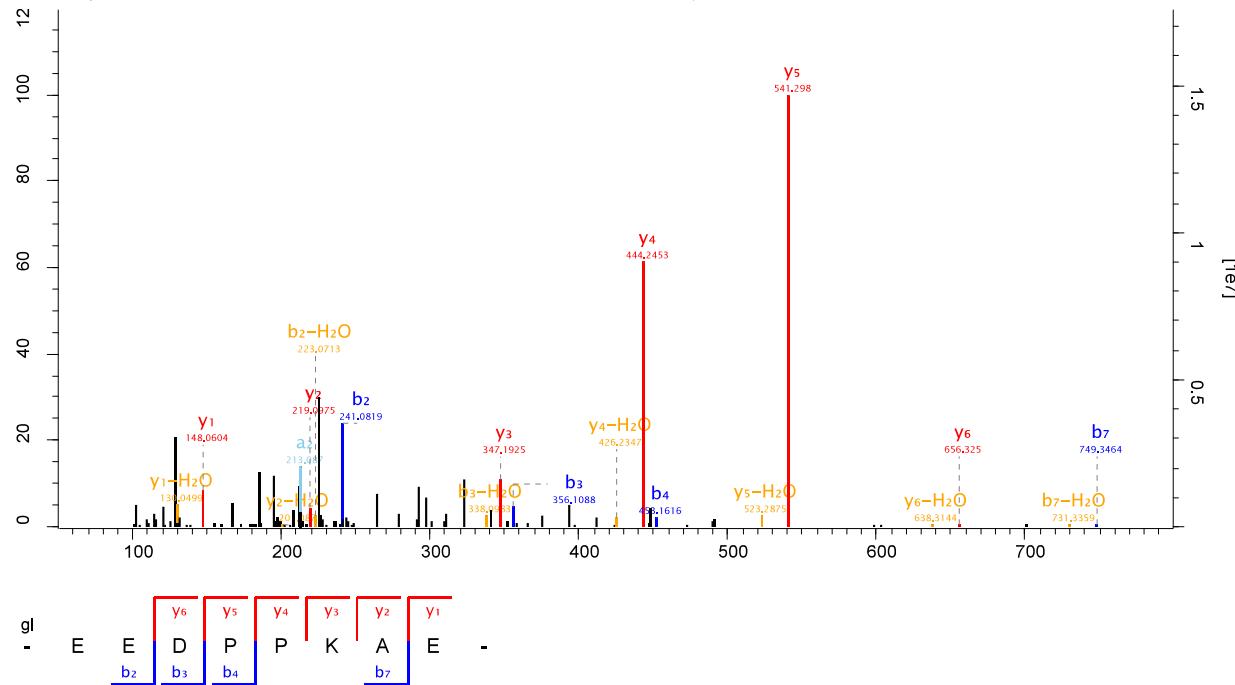
b₂ b₃ b₄ b₅ b₇ b₈ b₉ b₁₀ b₁₁ b₁₂

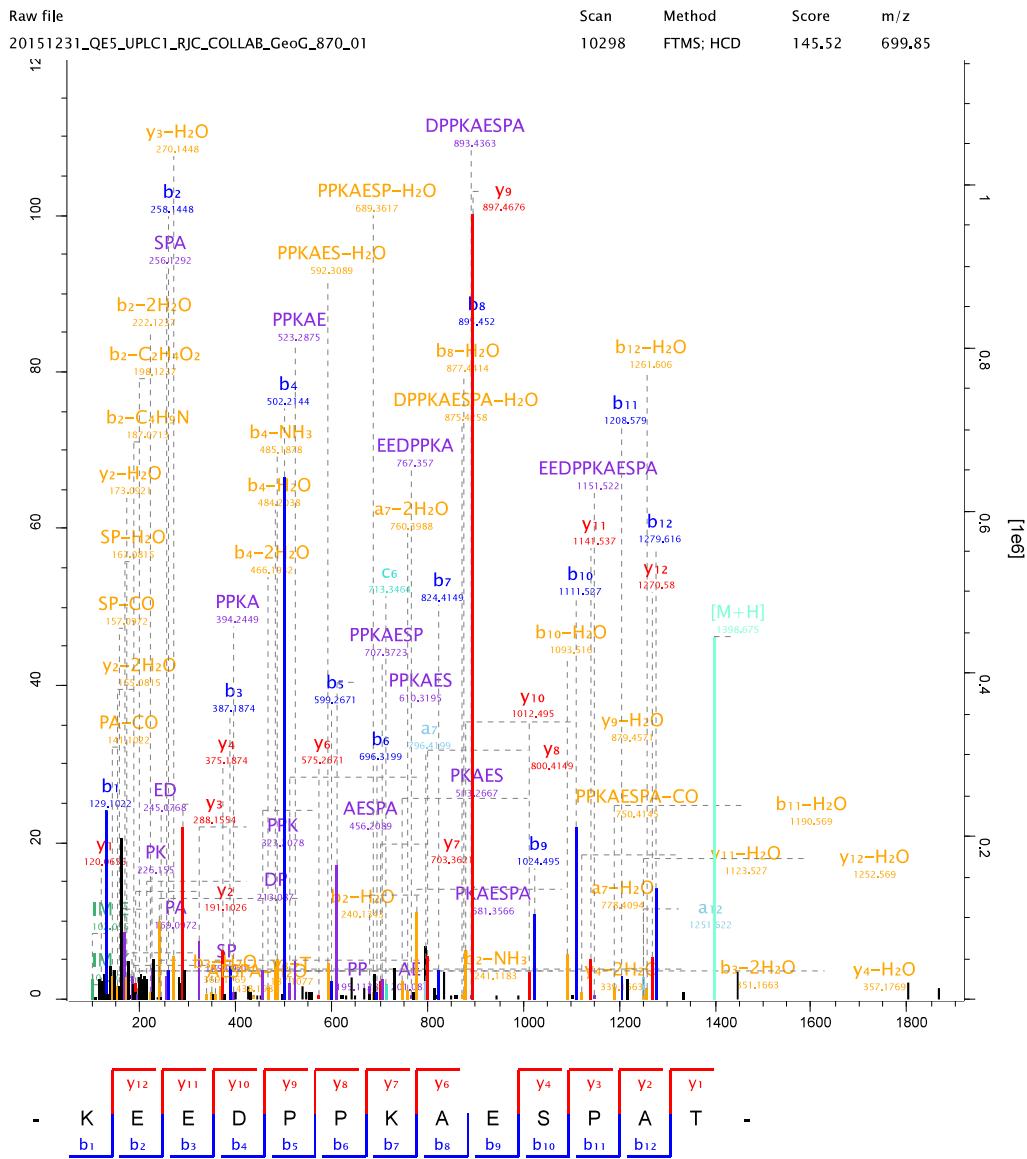
y₁₁ y₁₀ y₉ y₈ y₅ y₄ y₃ y₂ y₁



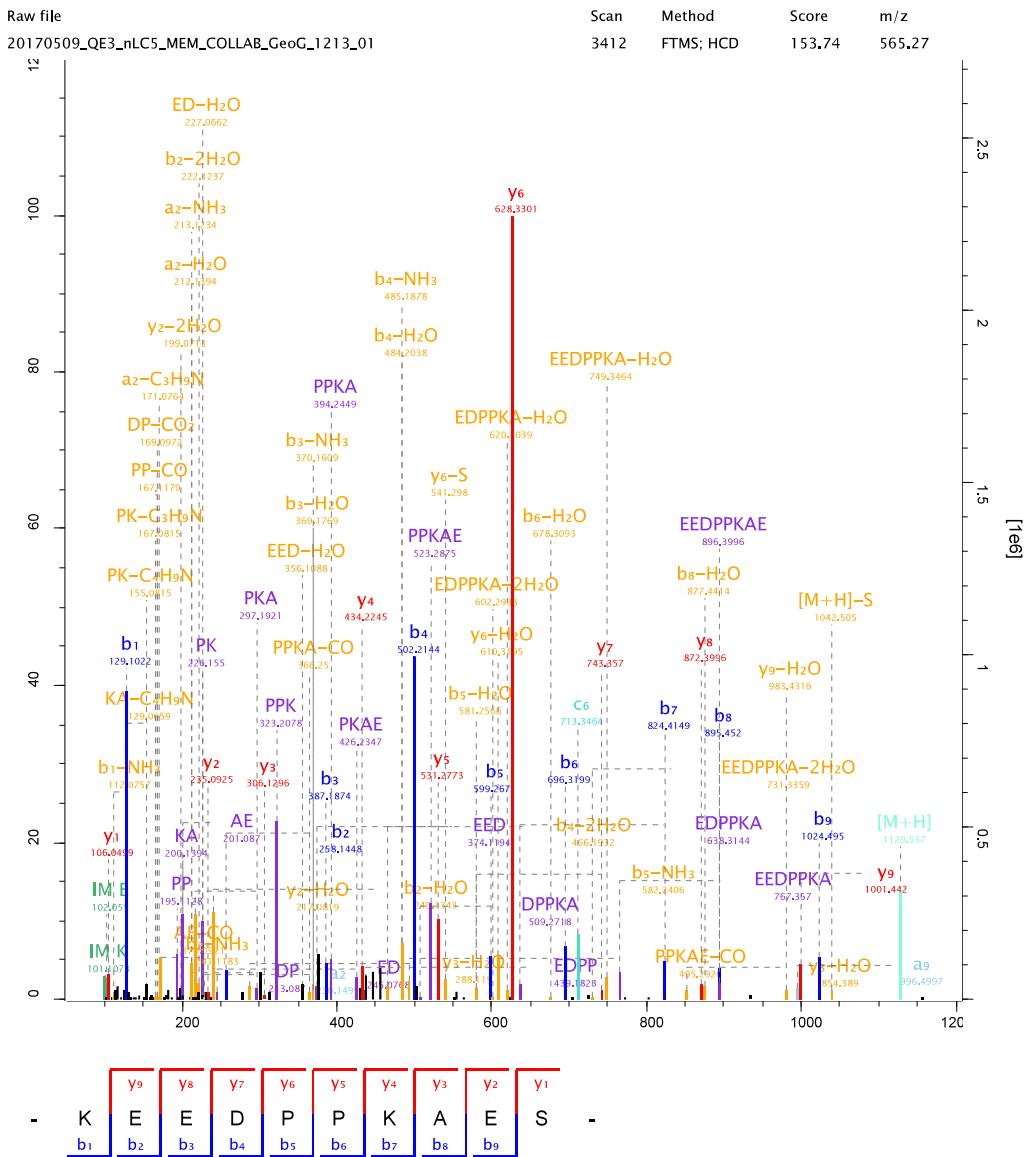
Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1213_01

Scan 9182 Method FTMS; HCD Score 101.25 m/z 449.2





K	E	E	D	P	P	K	A	E	S	P	A	T
b ₁	y ₁₂	y ₁₁	y ₁₀	y ₉	y ₈	y ₇	y ₆	y ₅	y ₄	y ₃	y ₂	y ₁

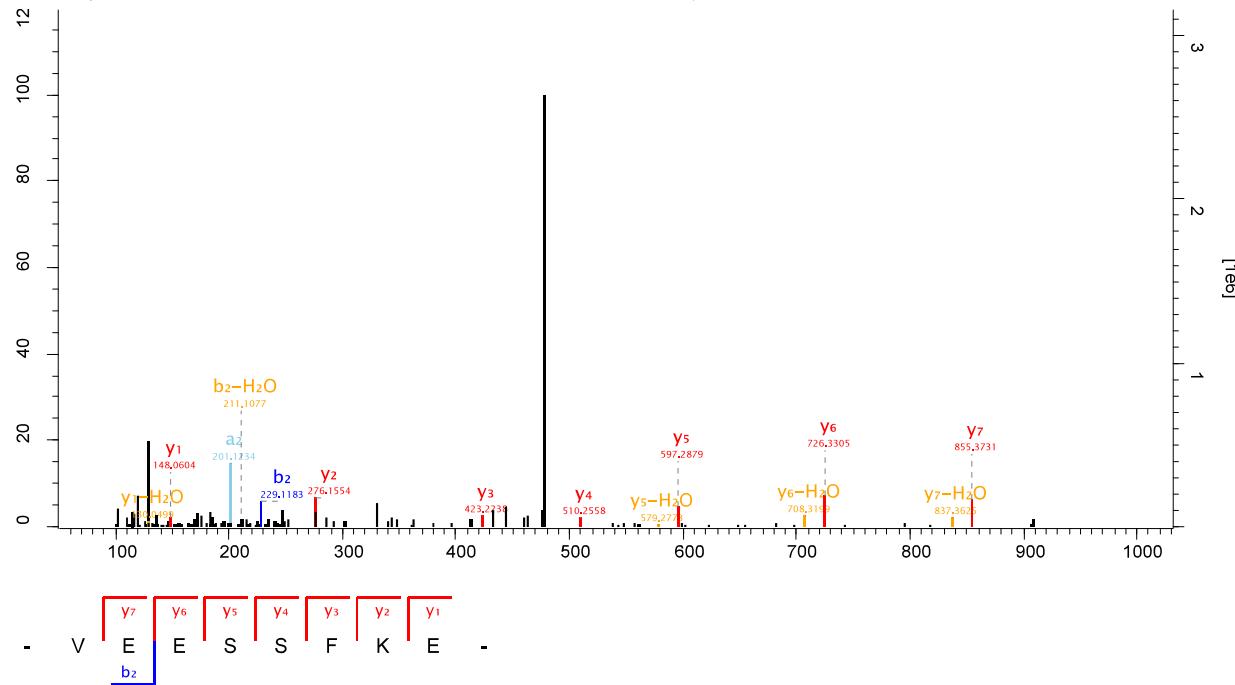


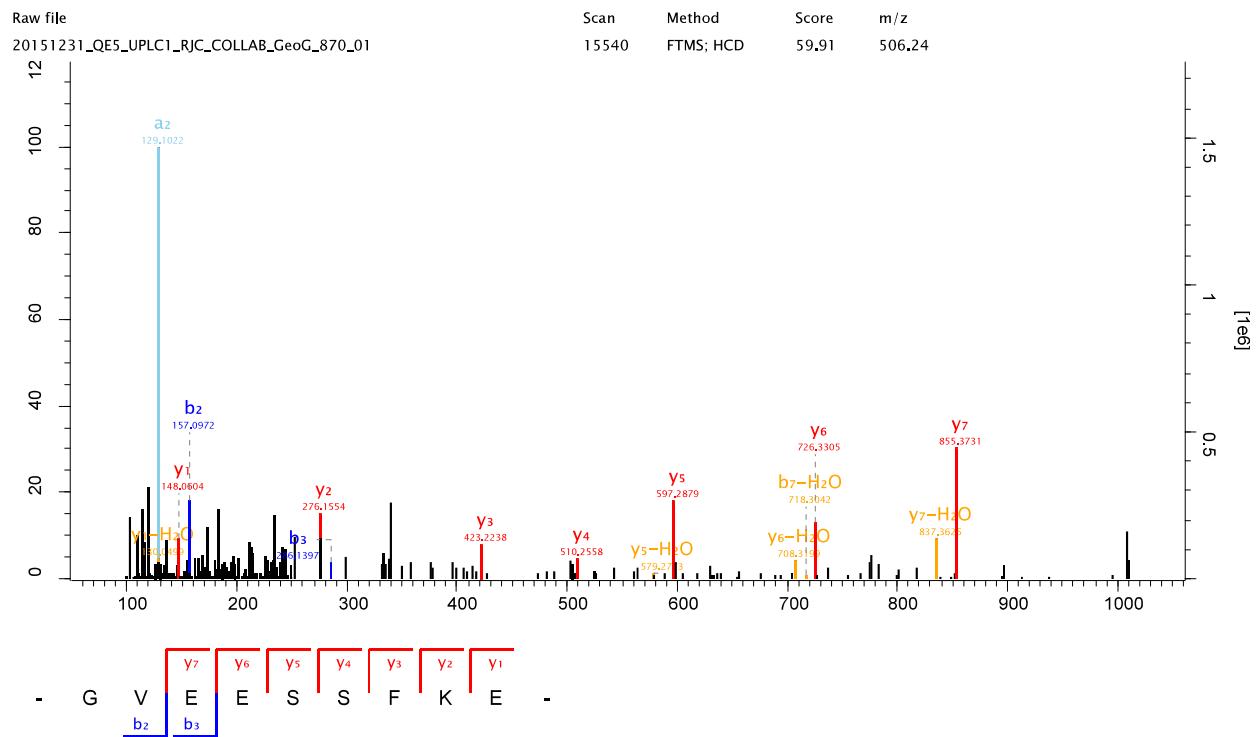
Sequence:

K	E	E	D	P	P	K	A	E	S
b ₁	y ₉	y ₈	y ₇	y ₆	y ₅	y ₄	y ₃	y ₂	y ₁

Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01

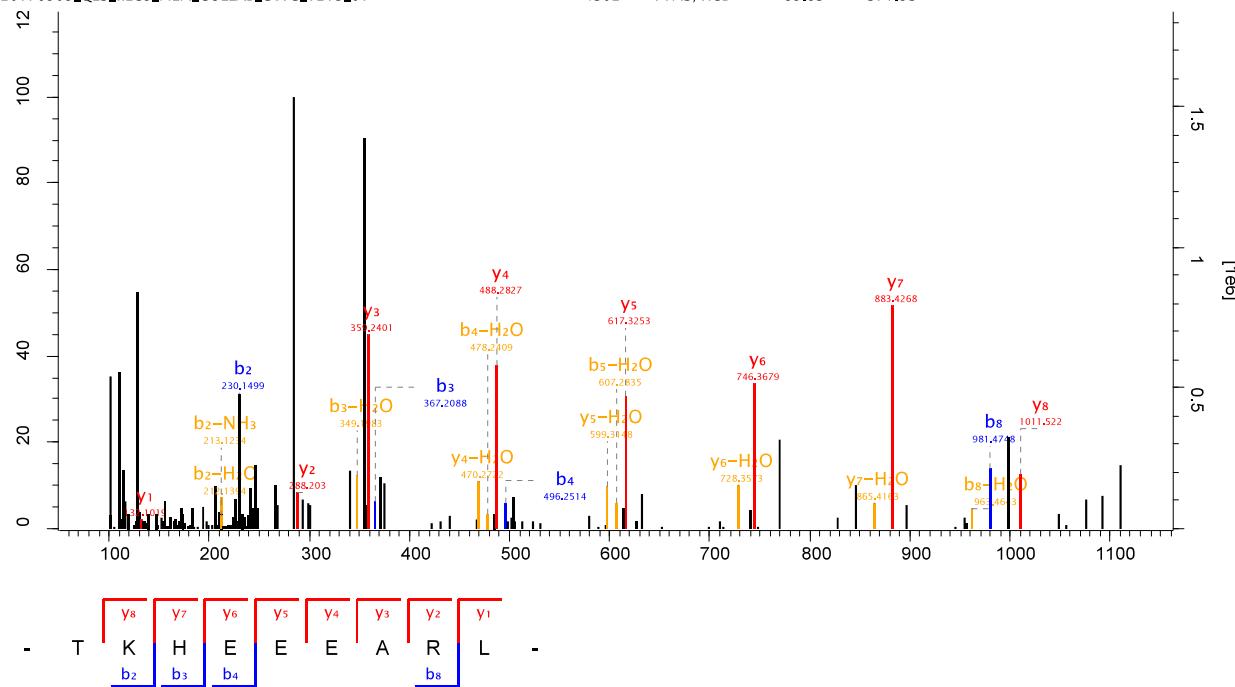
Scan 7459 Method FTMS; HCD Score 65.22 m/z 477.73

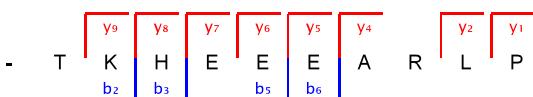
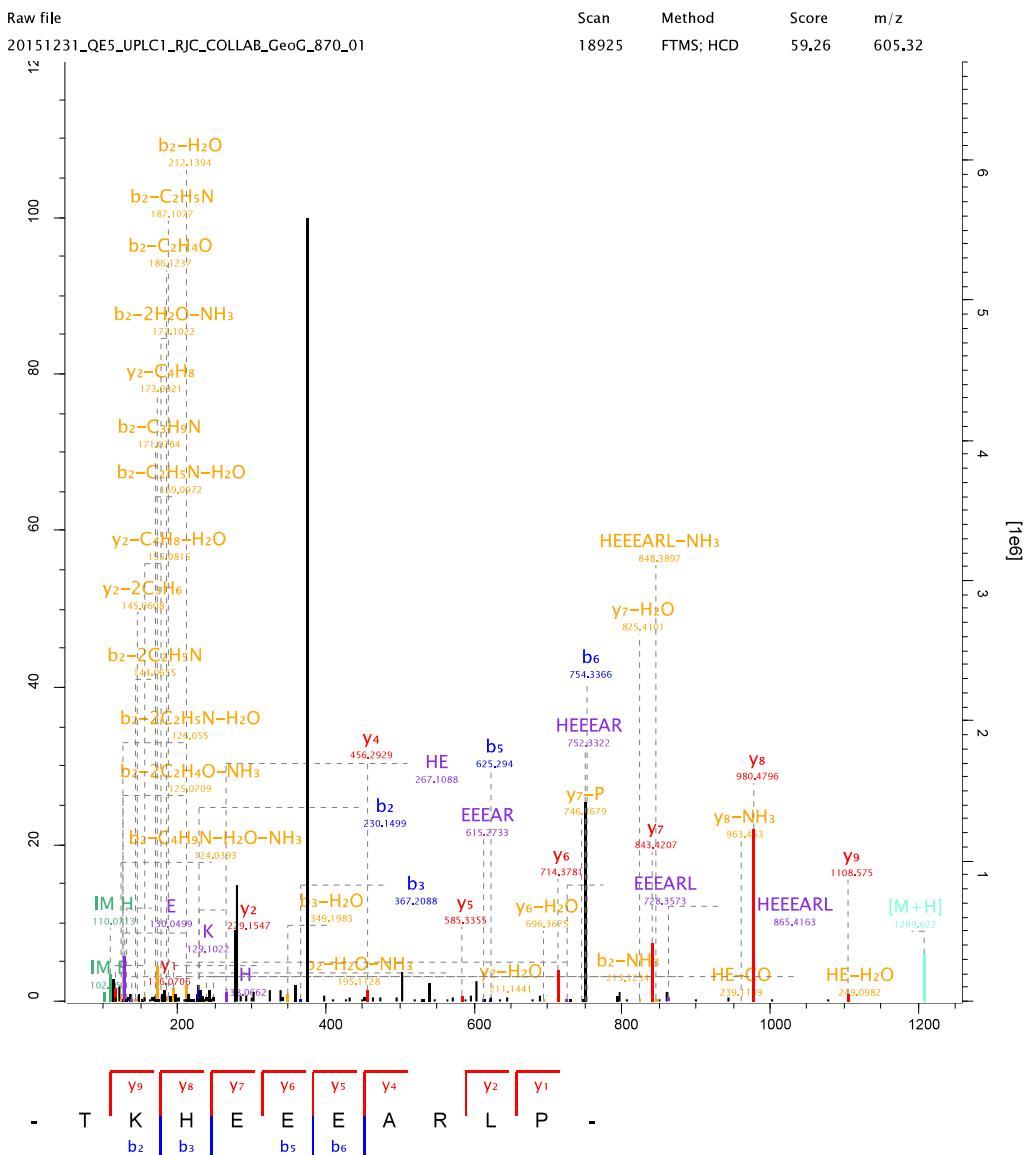


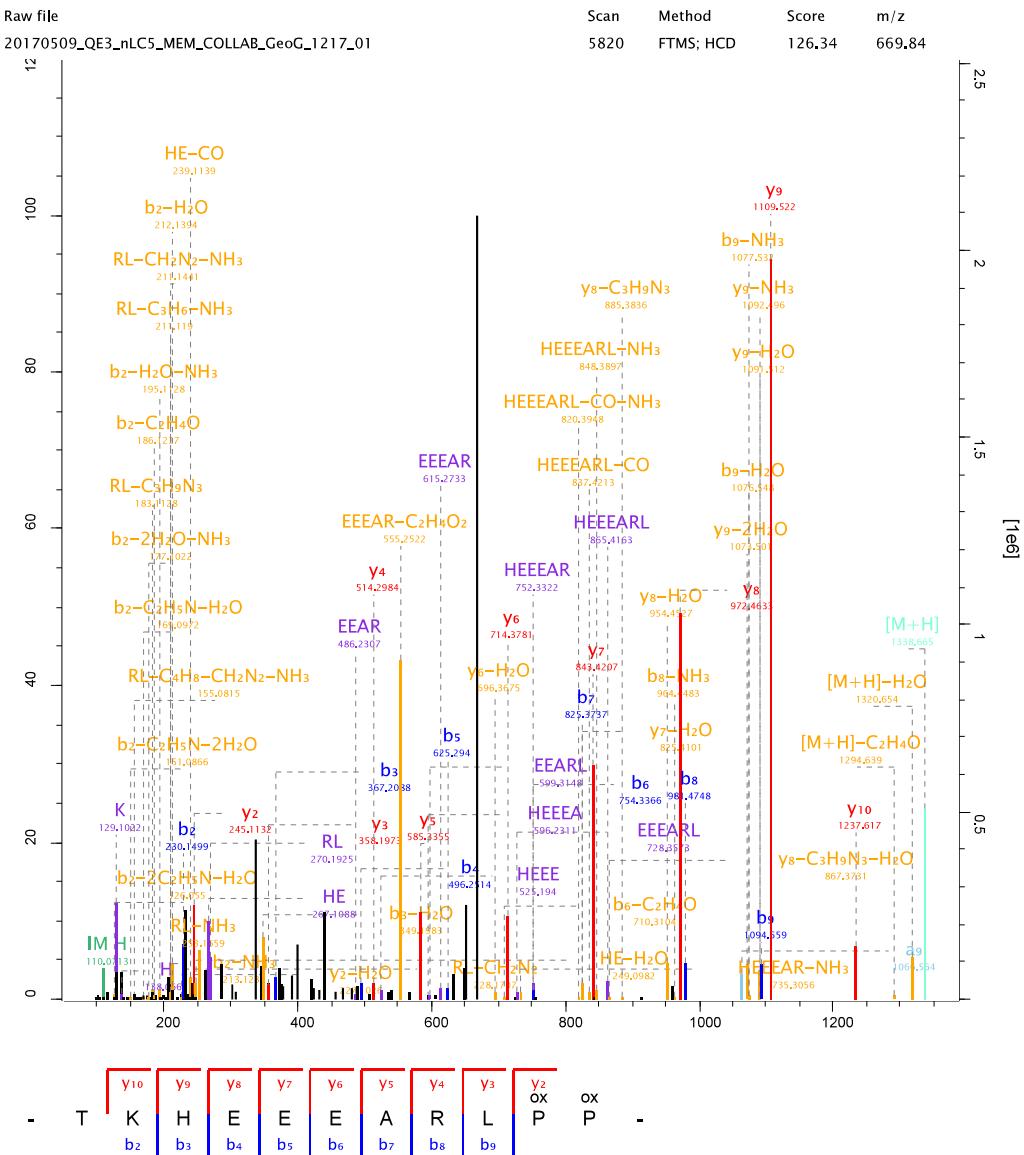


Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1213_01

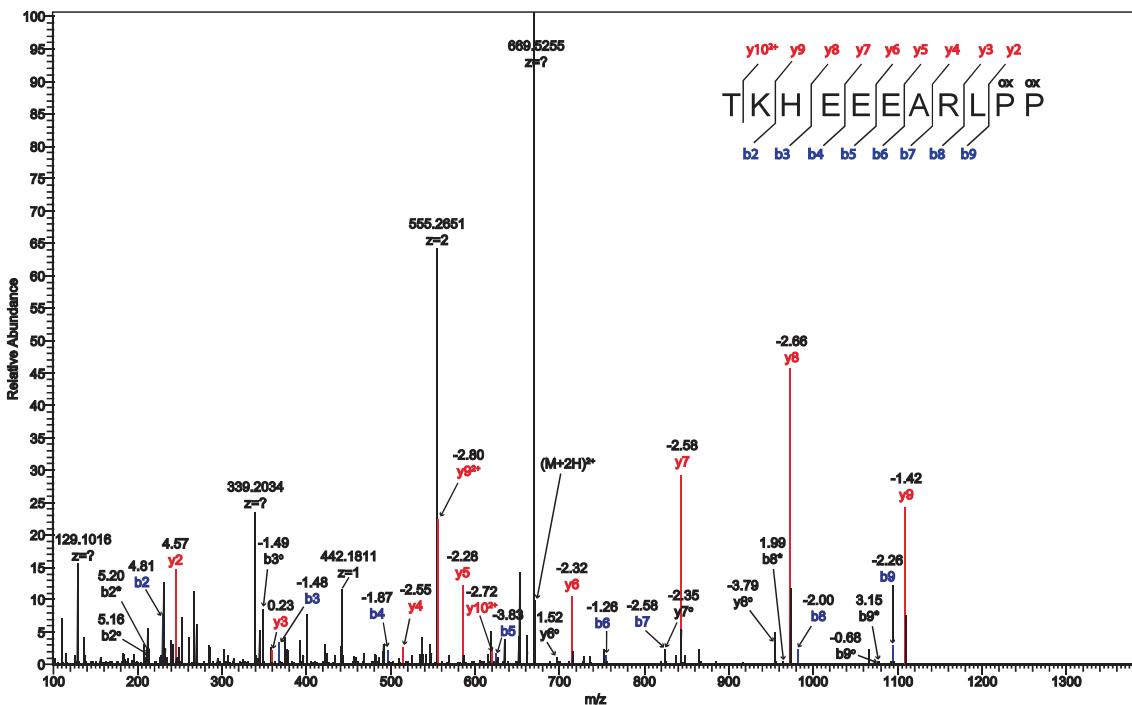
Scan 4362 Method FTMS; HCD Score 89.83 m/z 371.53

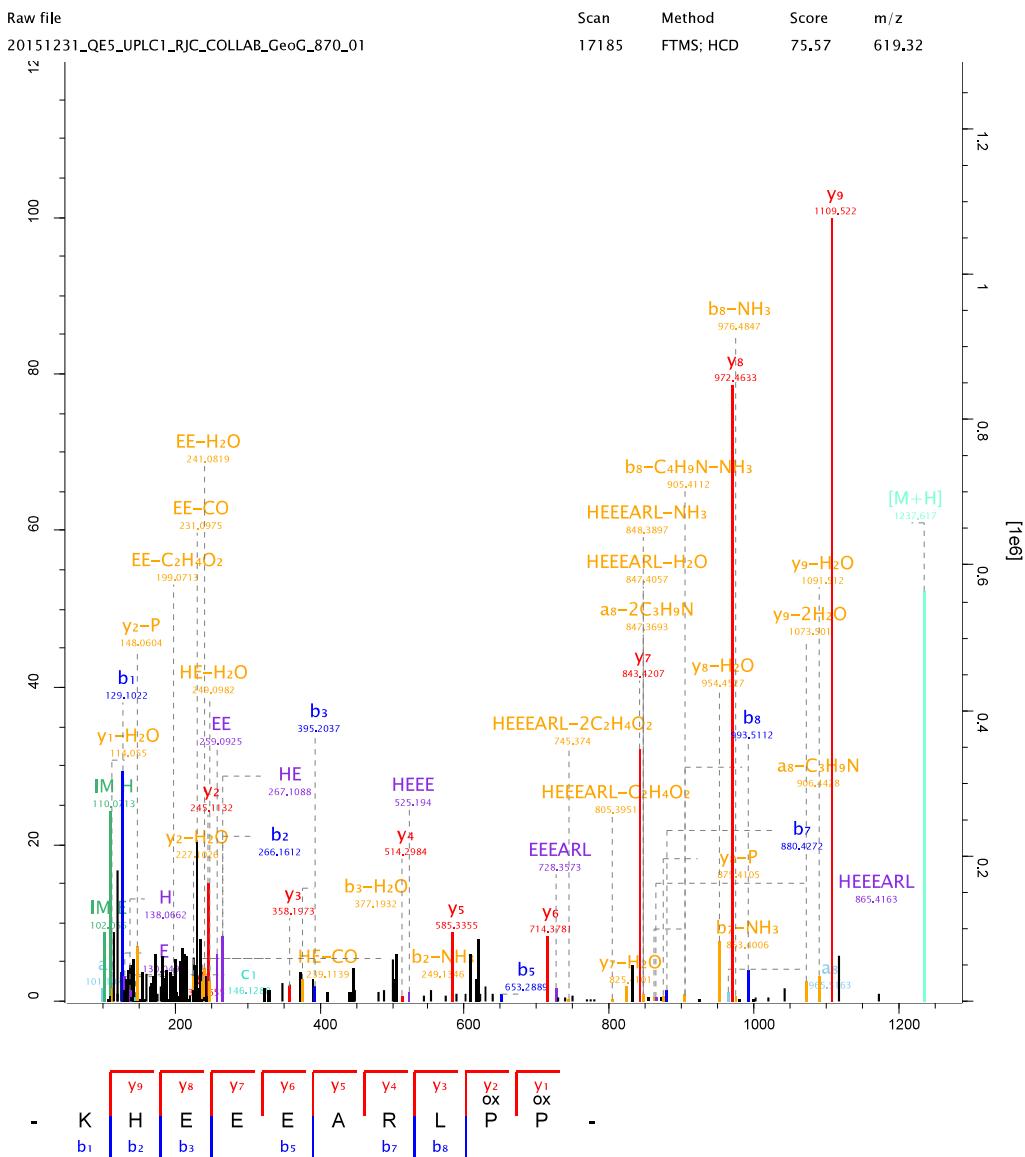




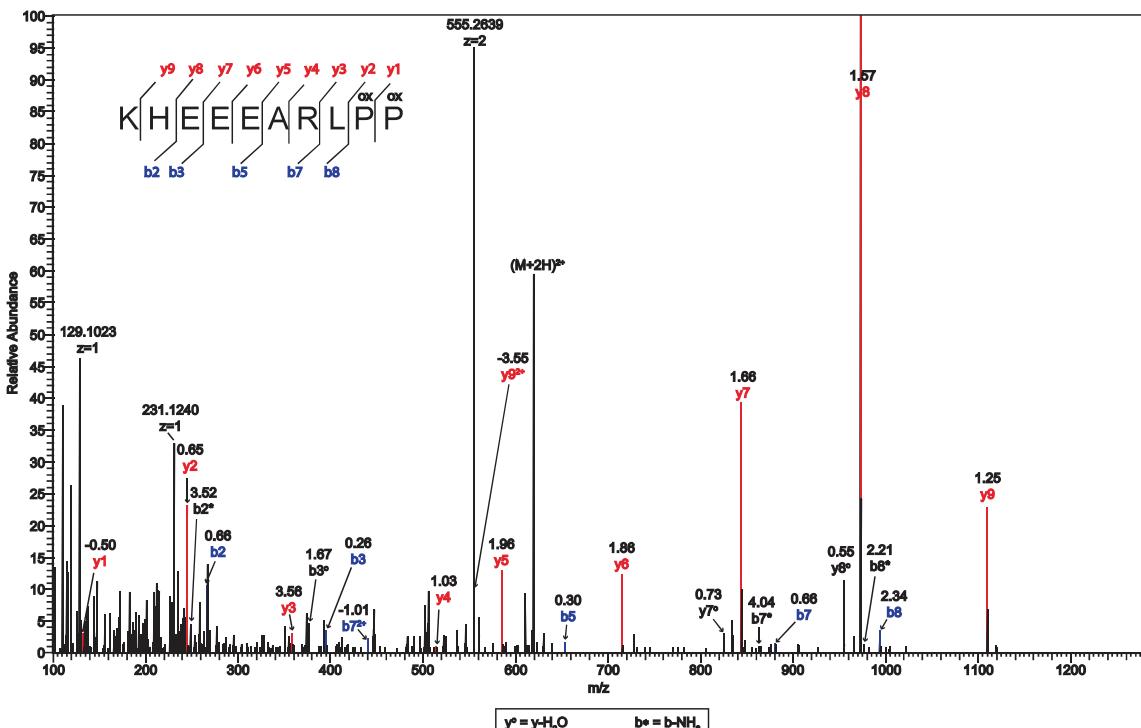


20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01#6820 RT: 22.17 AV: 1 NL: 4.63E5
T: FTMS + p NS1d Full ms2 689.84@hod28.00 [100.00-1385.00]



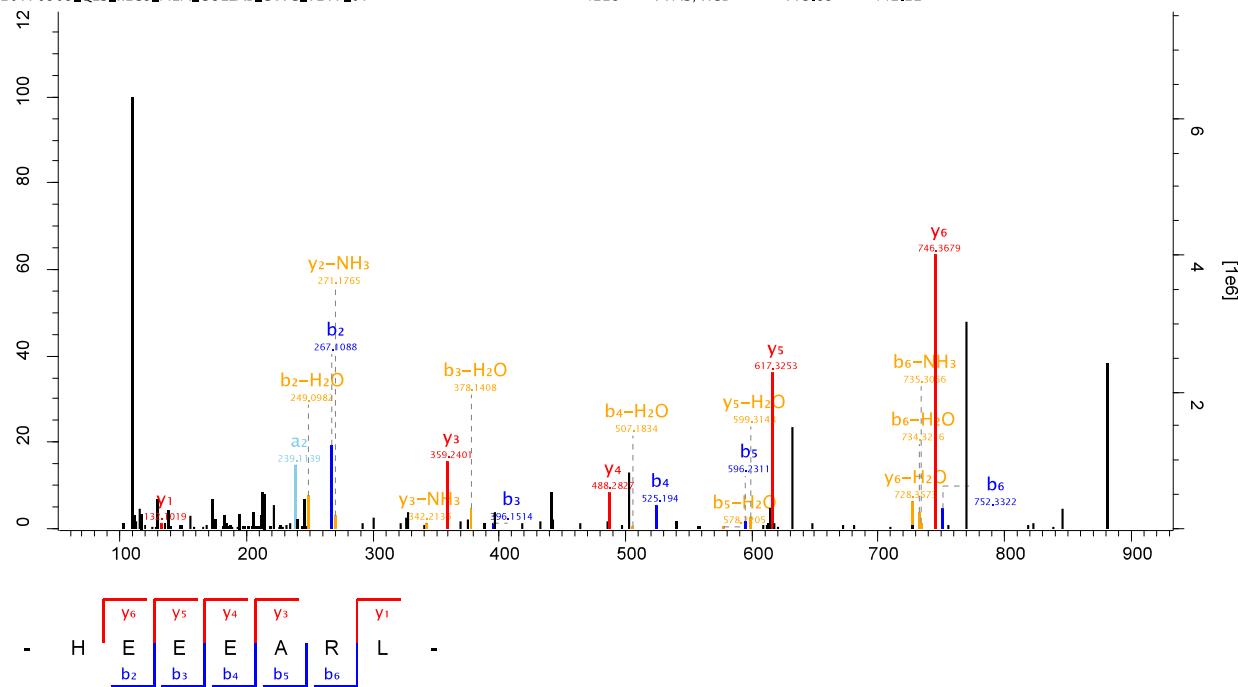


20151231_QE5_UPLC1_RJC_COLLAB_GeoG_870_01 #17185 RT: 61.55 AV: 1 NL: 1.73E5
T: FTMS + p NSId Full ms2 619.32@hcd28.00 [100.00-1280.00]



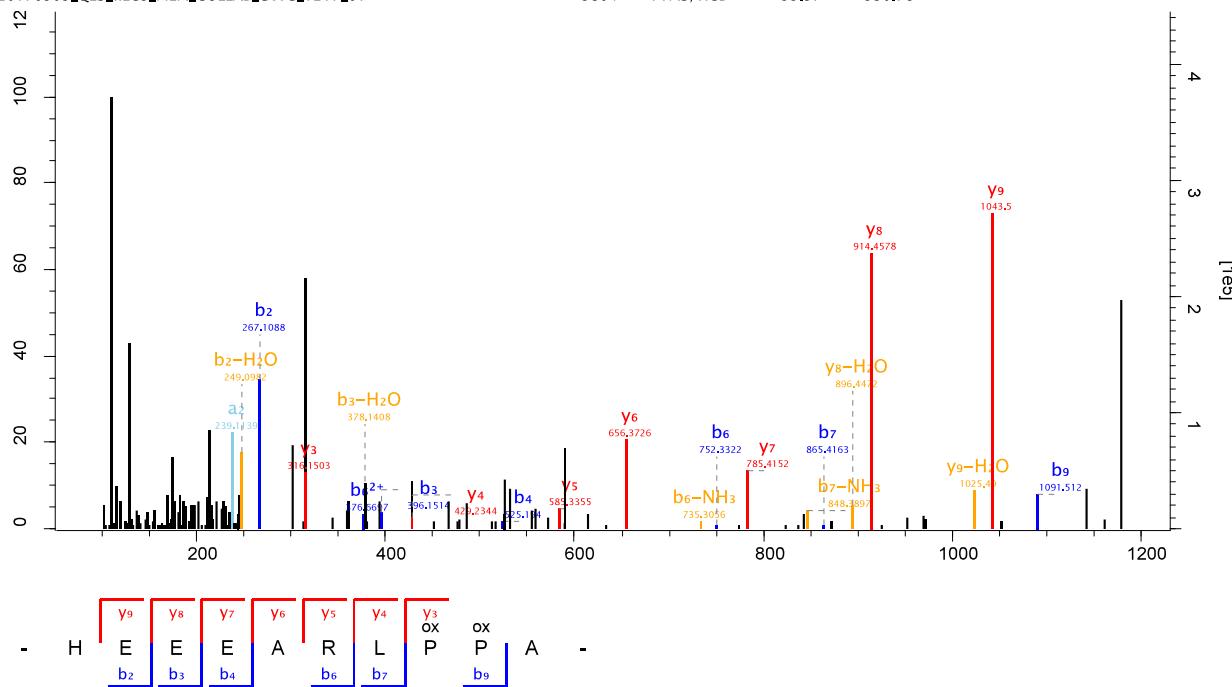
Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01

Scan 4229 Method FTMS; HCD Score 113.89 m/z 442.22

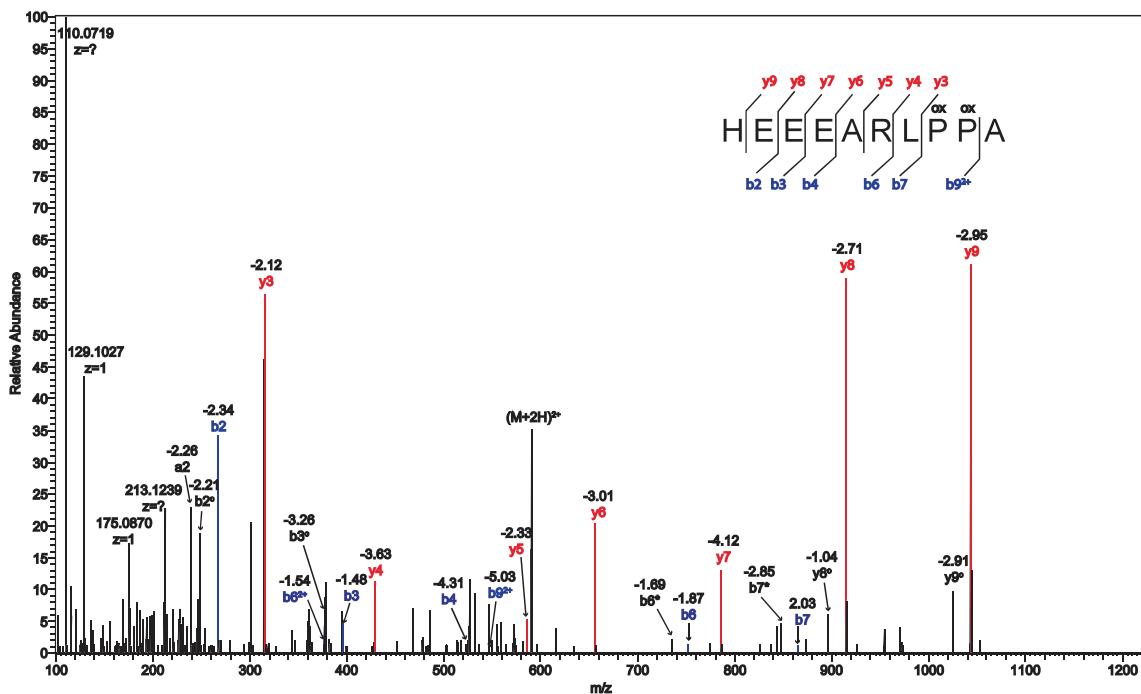


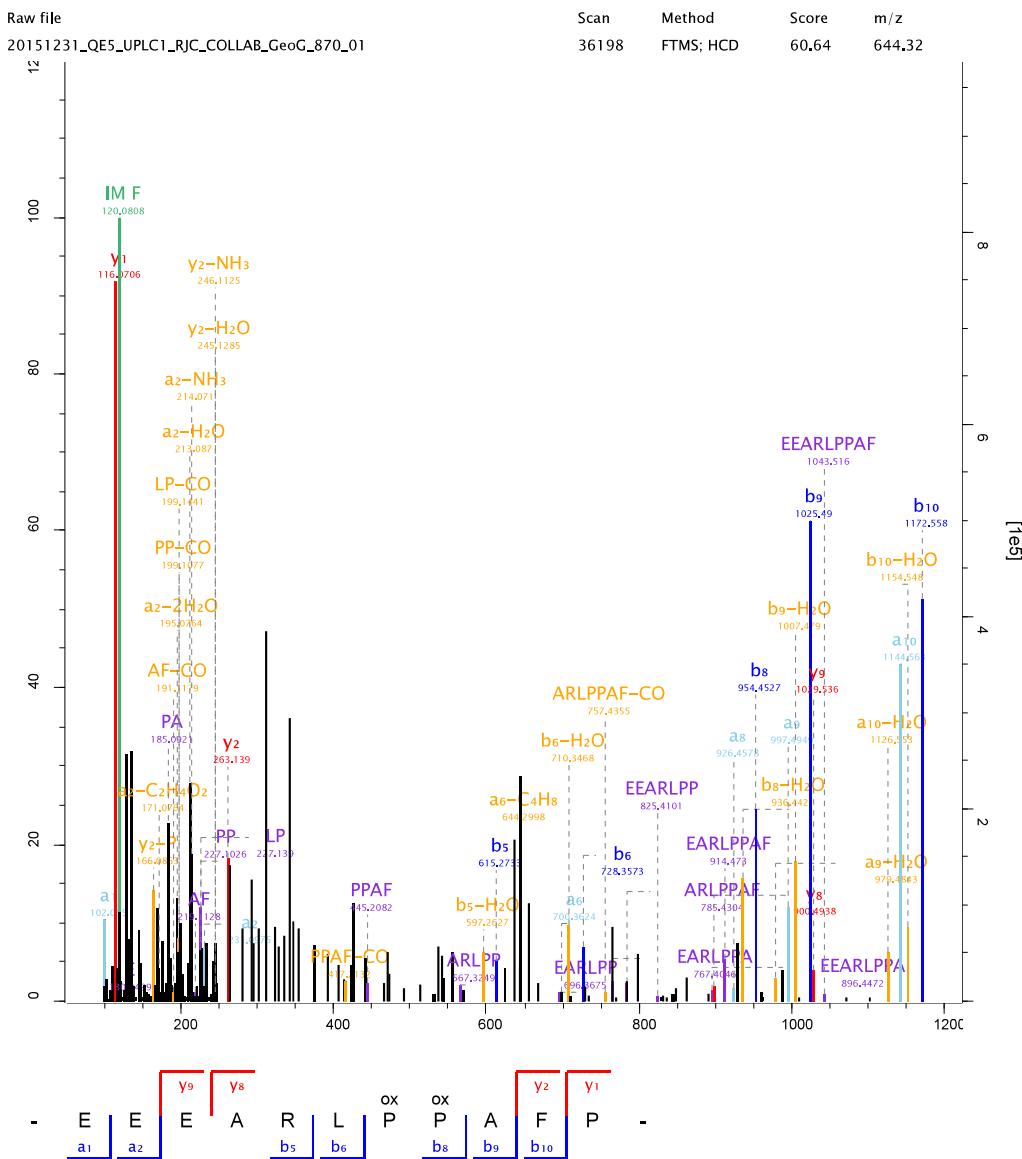
Raw file
20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01

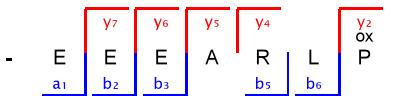
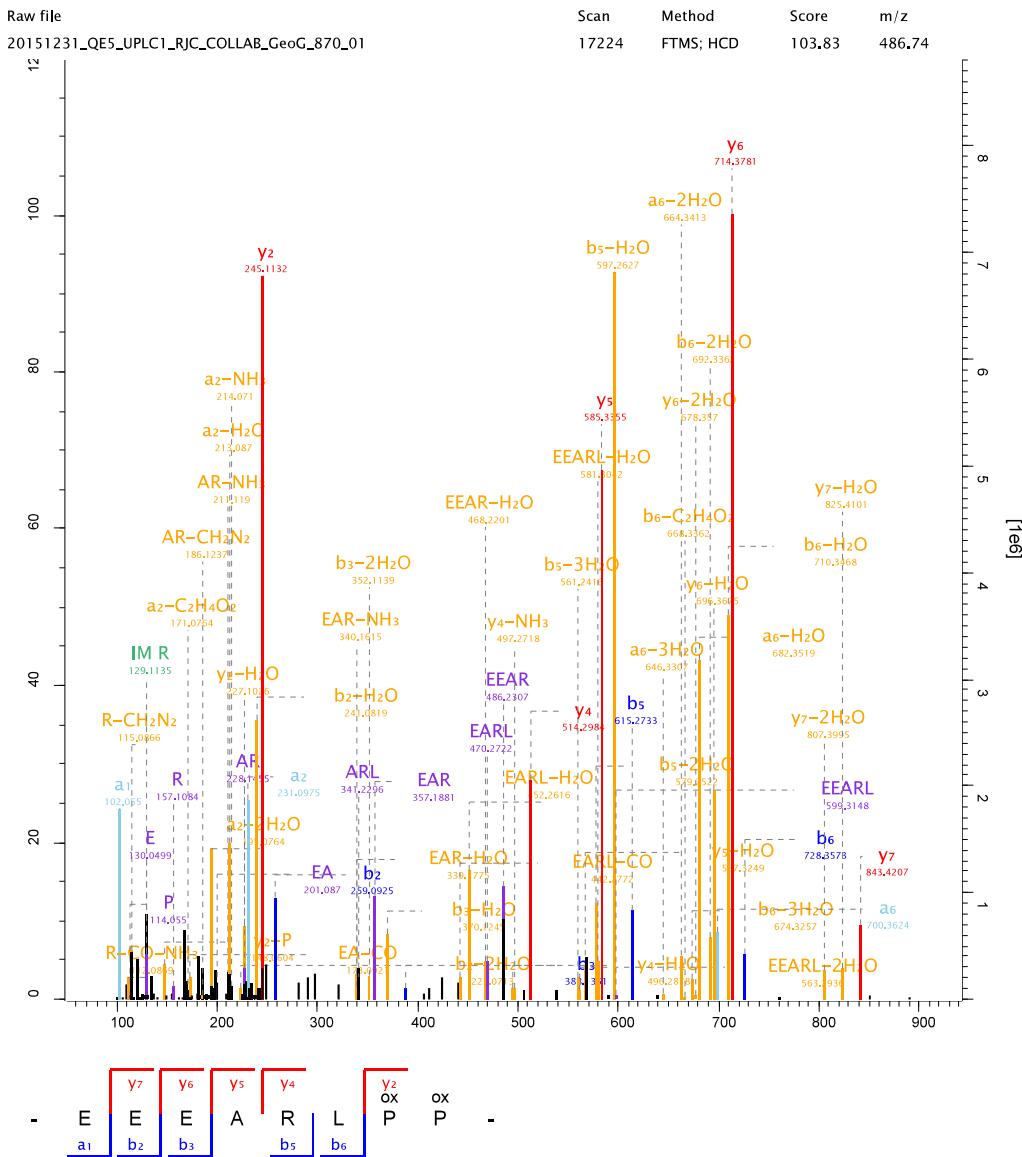
Scan 9804 Method FTMS; HCD Score 99.57 m/z 590.78

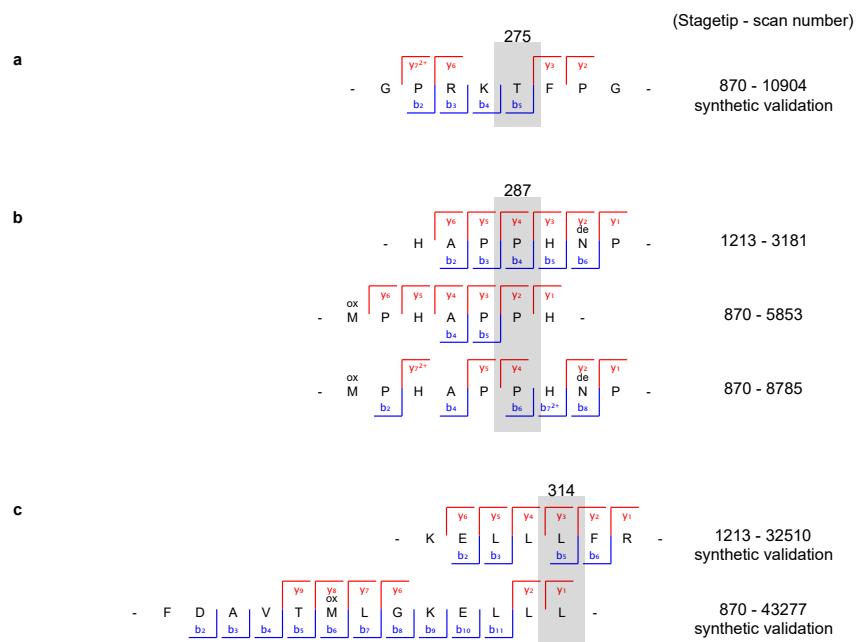


20170509_QE3_nLC5_MEM_COLLAB_GeoG_1217_01 #9804 RT: 34.30 AV: 1 NL: 9.18E4
T: FTMS + p NSI d Full ms2 590.78@mz28.00 [100.00-1225.00]

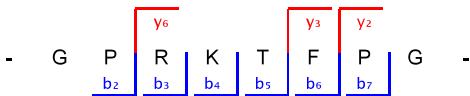
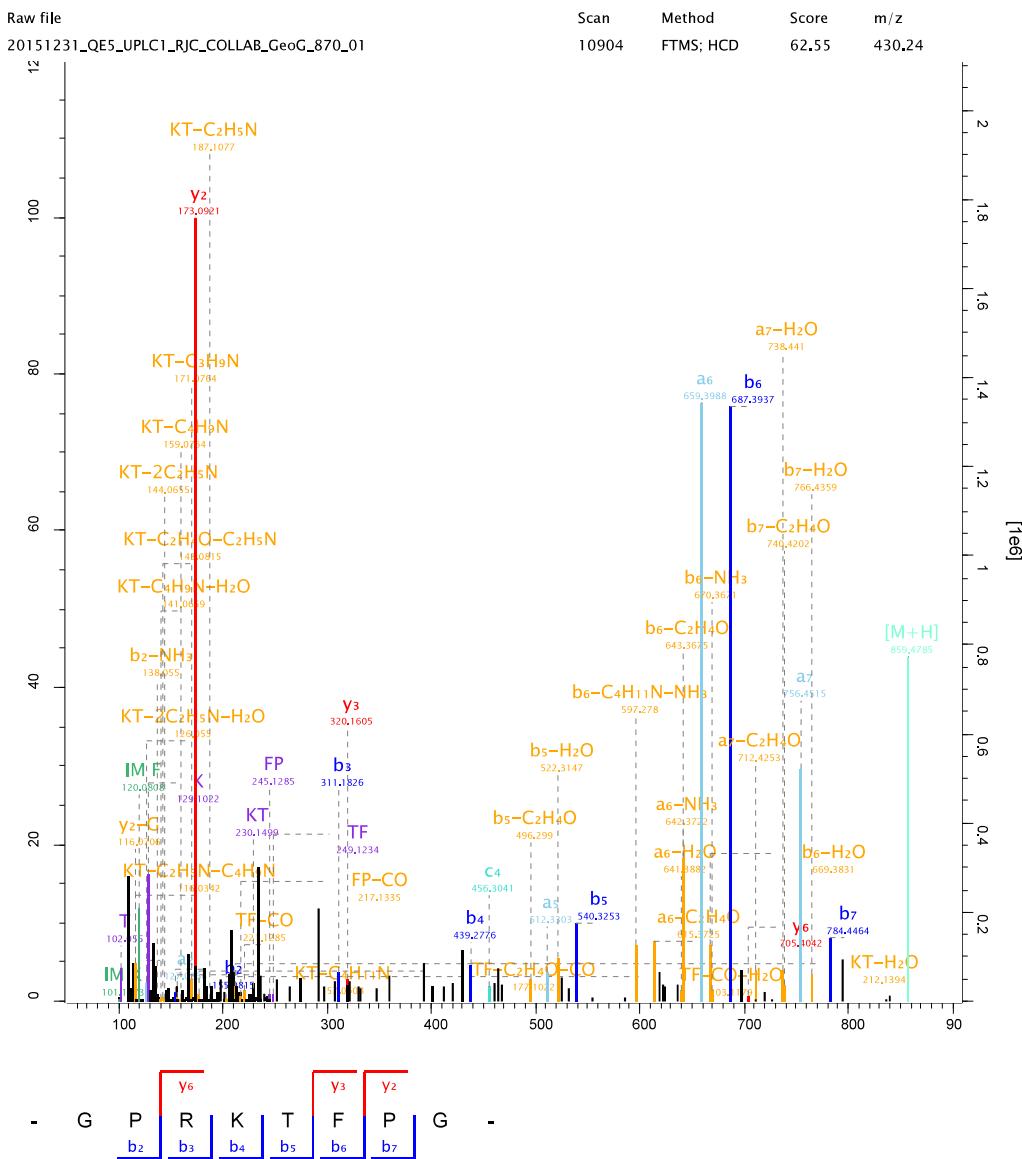


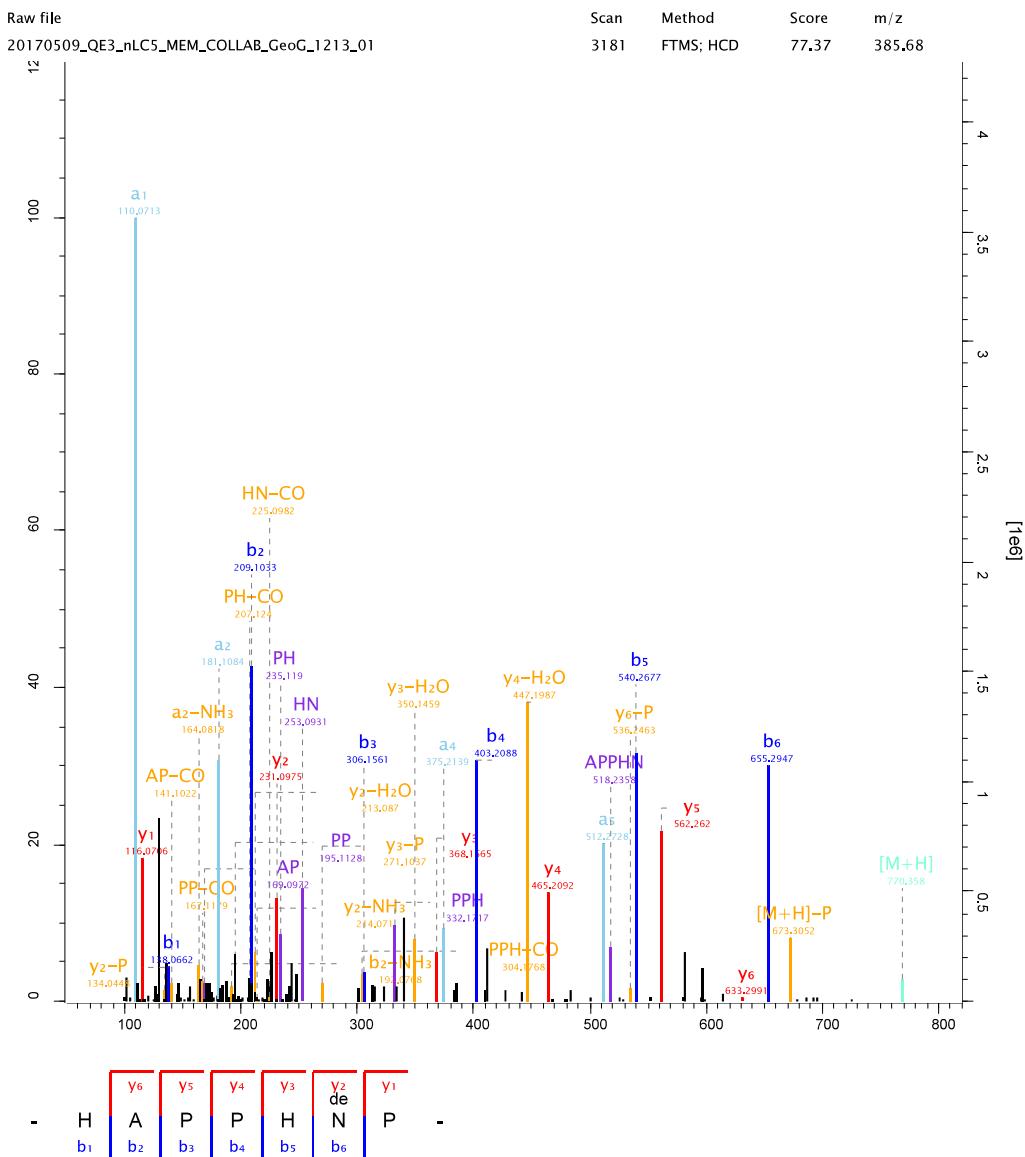


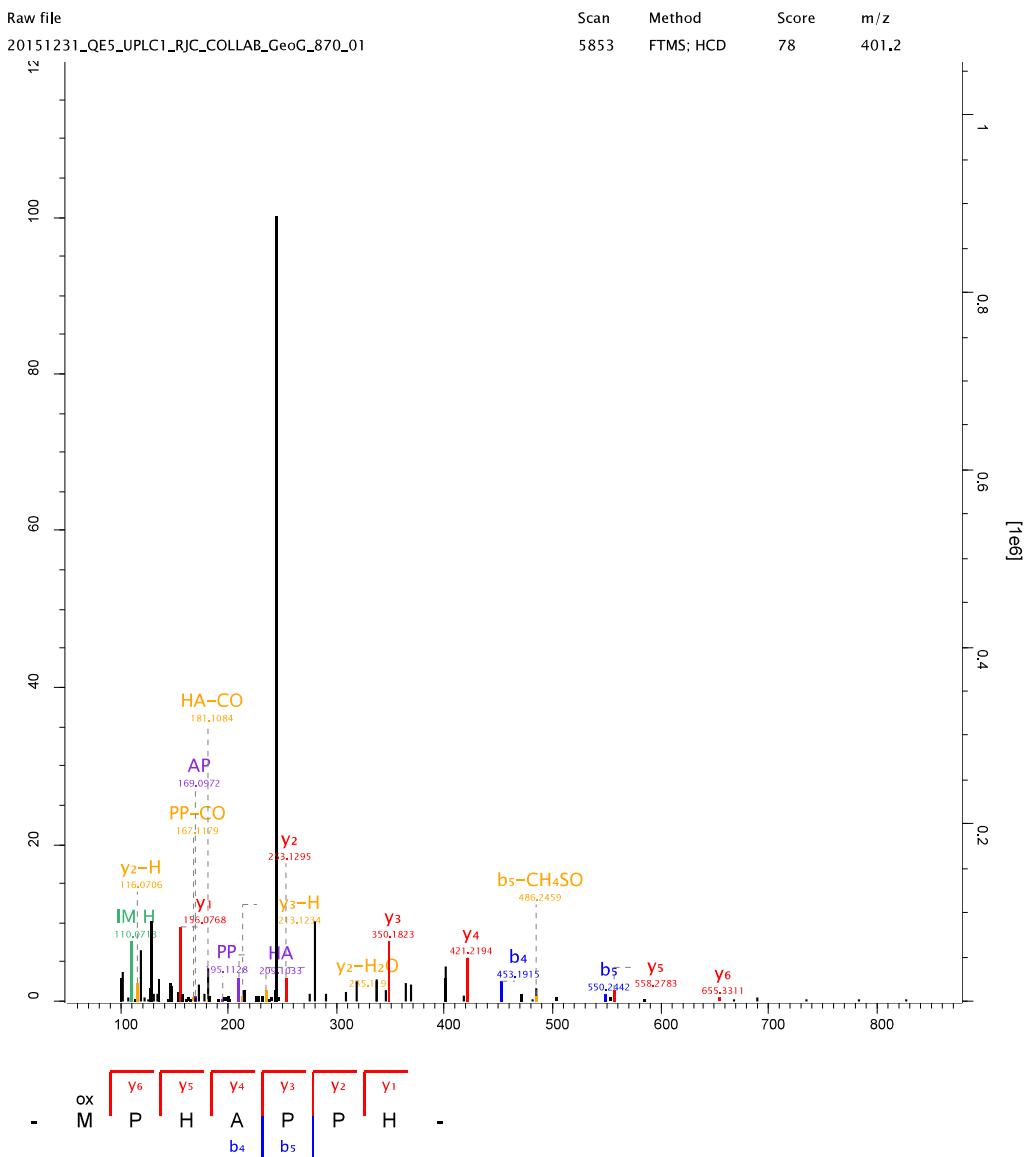


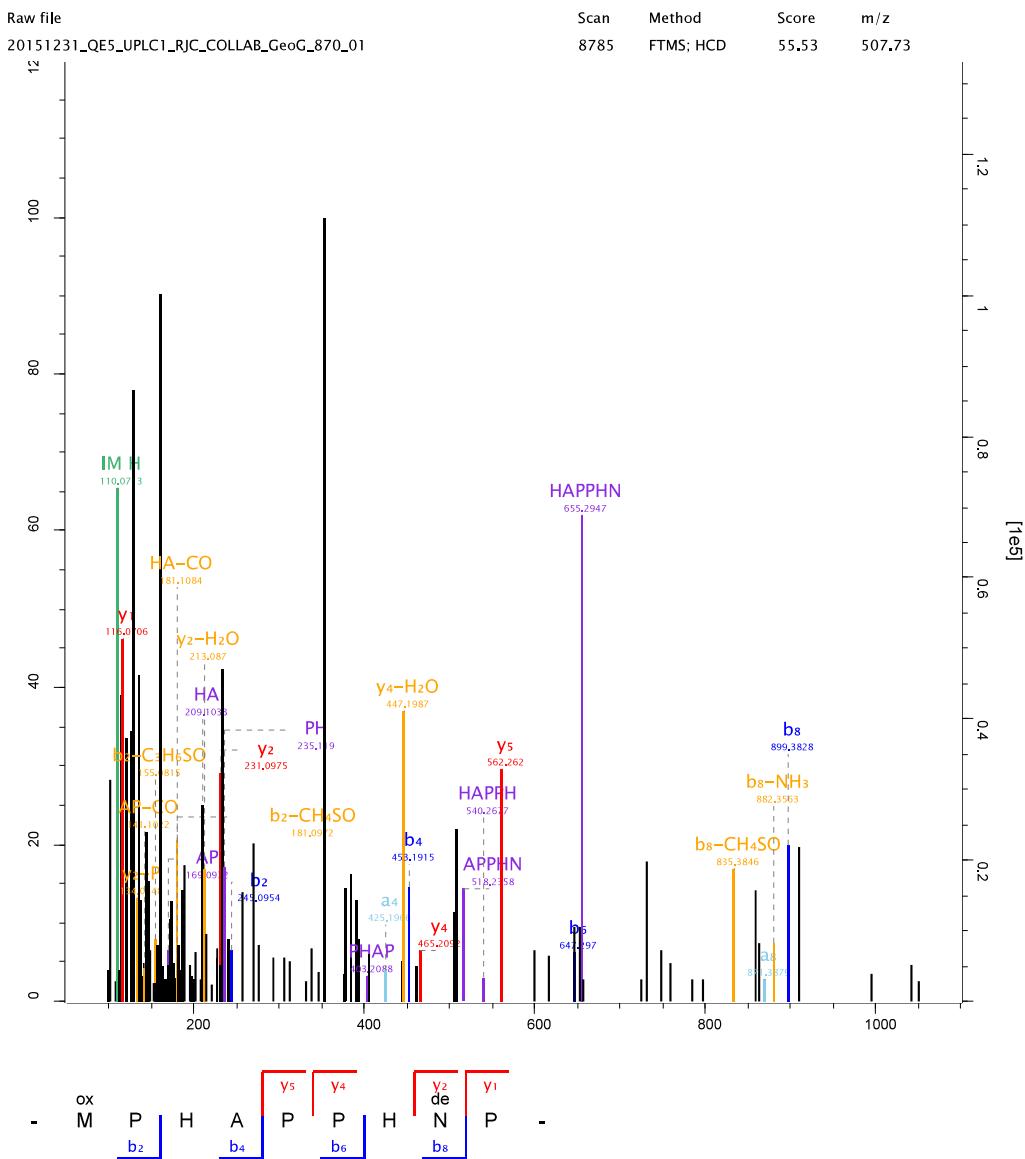


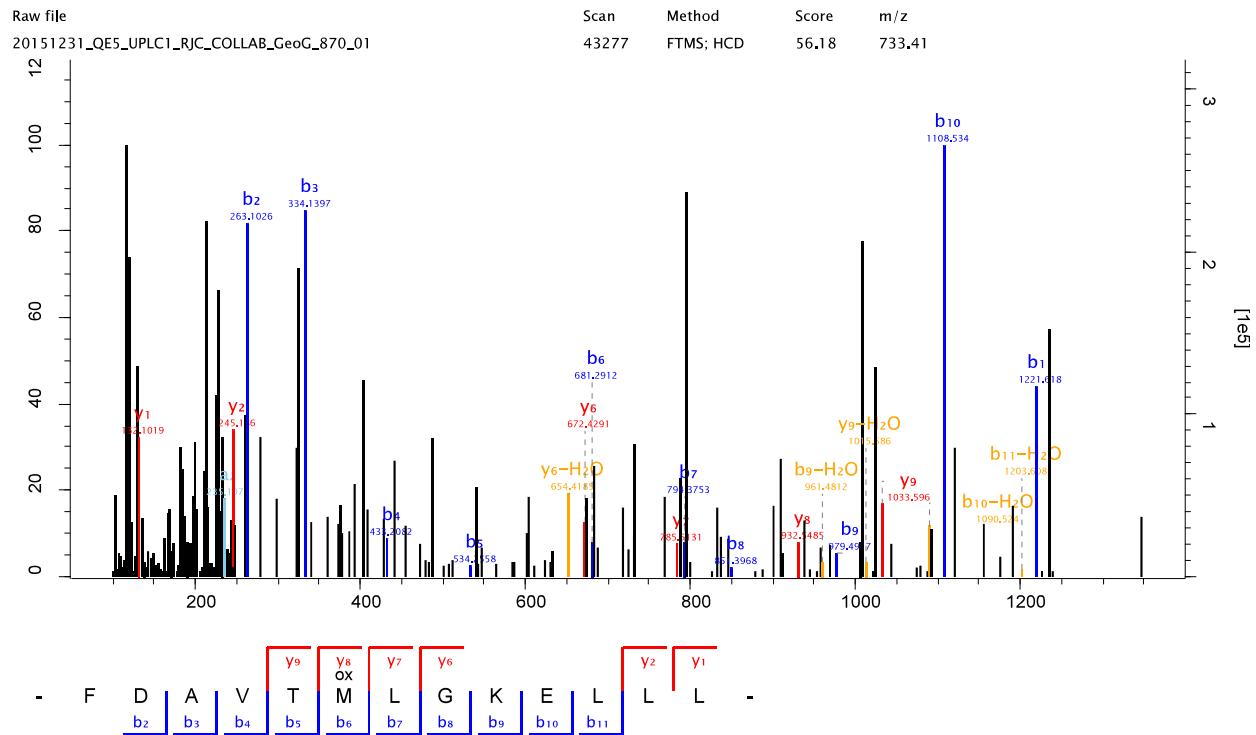
MMP20. Peptide matches to phylogenetically informative positions 275 (a), 287 (b), and 314 (c). Coordinates in reference to XP_004427457.

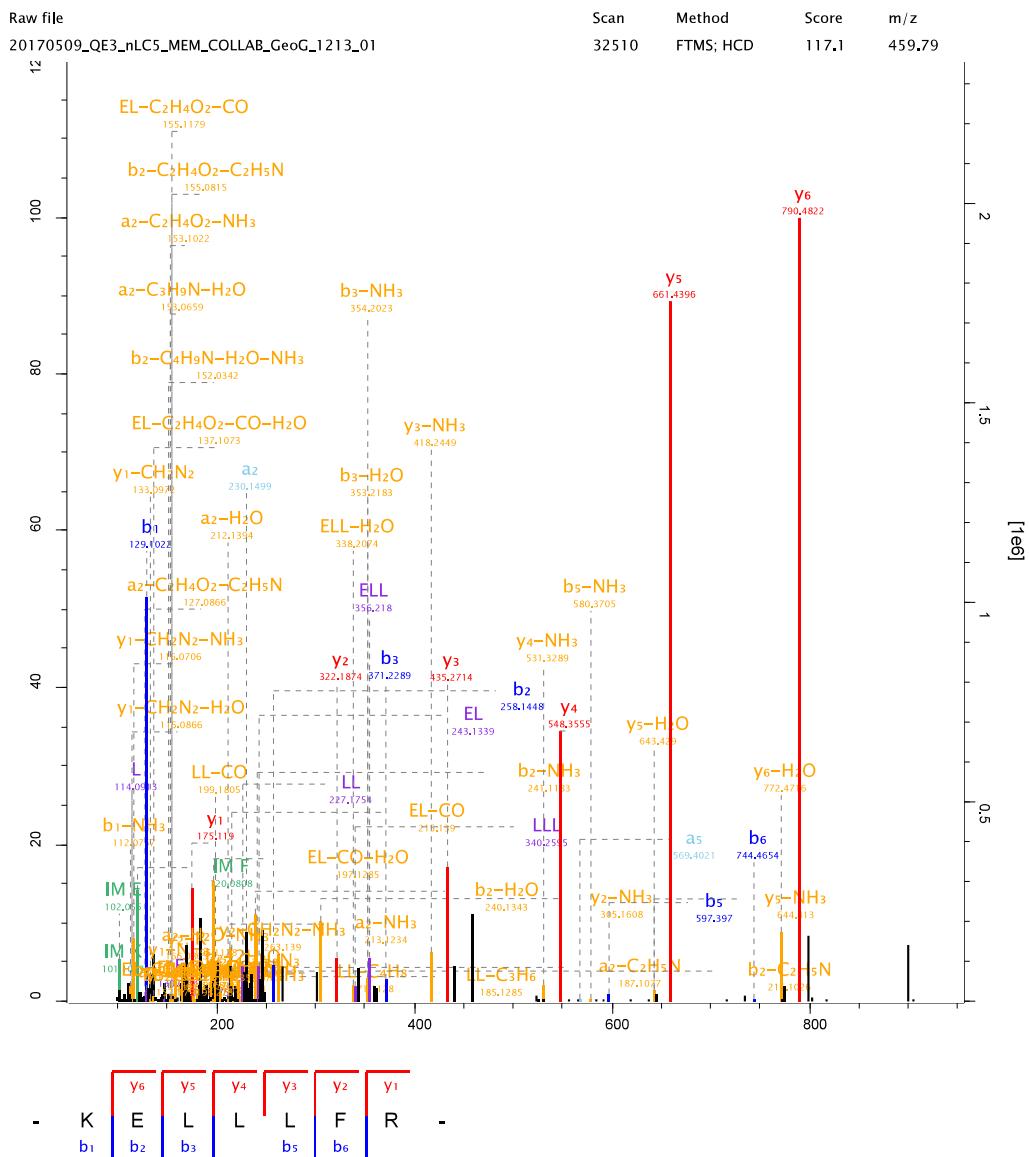












Synthetic Peptides

The following pages contain the manual annotation of phylogenetically informative peptides ($n=18$) and phosphopeptides ($n=3$) confirmed through synthetic production of the peptide sequences identified through MaxQuant. Both synthetic and ancient peptides are manually annotated through the Interactive Peptide Spectral Annotator (IPSA). These annotations, and associated mass errors, indicate almost complete correspondence between the two peptide datasets. This includes the phosphopeptides.

For each peptide, the MS2 spectrum of the synthetic peptide is shown at the top of the page, and the corresponding MS2 spectrum of the ancient peptide is shown at the bottom of the page. MaxQuant fragment ion annotation (automated) for the latter spectra can be found earlier in the document.

Note that some spectra contain additional high-intensity fragment ions that are not identified as y- or b-ions. This is likely due to the cofragmentation of impurities present in the synthetic and/or ancient peptide mixture.

